Application Number 10/772,537
Art Unit 1639

REMARKS

This response is filed to place the above-referenced case in condition for immediate allowance. As to the claims, claims 1 through 10 were previously cancelled. Applicant further cancels claims 12 through 14 and adds new claims 16 through 19, antecedent basis for which may be found in the specification on page 43 lines 5 through 17. Applicant cancels claim 13 but plans to file a divisional application based on the decapeptide. Applicant further amends claims 11 and 15 and the specification on pages 4, 40, 43 and the Abstract to comply with the sequence rules. No new matter has been added. Reexamination and reconsideration of the application, as amended, is requested.

Response to Improper Sequence Identifications

Examiner has noted that Applicant has improperly identified its amino acid sequences. Specifically, Examiner has stated, "applicant's specification and especially the claims MUST be amended to insert proper SEQ. ID. to satisfy the sequence rules." The sequence rules require the use of three-letter codes for amino acids as set forth in the tables of WIPO Standard DT.25 (1998). *See* 37 CFR 1.822; *See also* MPEP ¶¶ 2422 and 2423. Applicant has respectfully amended the specification and the claims so that they list the proper SEQ. ID. in compliance with the sequence rules.

Response to Restriction/Election Requirement

Examiner has subjected claims 11 to 15 to a restriction/ election requirement. As to claims 11 to 15, Examiner has stated that:

> "claims 11-15 [are] directed to the following patentably distinct species of the claimed invention: compounds (and their analogs) that comprise

> decapeptides of different amino acid sequence which require different and separately burdensome manual/ computer structure and bibliographic searches in patent and literature databases."

Furthermore, Examiner has required Applicant to elect a single disclosed species (e.g. a single decapeptide) under 35 U.S.C. 121 for prosecution on the merits.

Applicant respectfully traverses the requirement for election between the decapeptides pursuant to 37 CFR § 1.143 and provisionally elects the decapeptide of Claim 15 for prosecution on the merits. The restriction requirement is improper because claims 12, 14 and 15 are not "independent and distinct" from each other. Rather, the decapeptides are equivalent and differ by merely two amino acids, the second and fifth in sequence, neither of which affects the function of the whole sequence. The second and fifth amino acids are merely filler or generic amino acids and their respective space may be filled by any amino acid. Therefore, because the decapeptides in claims 12, 14 and 15 are in fact equivalent and do not require different and separately burdensome manual and computer searches by Examiner, the restriction requirement is improper.

To clarify the generic classification of the second and fifth amino acids in the decapeptides as disclosed in claims 12, 14 and 15, Applicant has respectfully submitted new claims 16 through 19. In independent claim 16, Applicant has denoted those generic amino acids with X02 and X02, which can independently be any amino acid. Persons skilled in the art would follow that the decapeptides are parcel of the same concept and that the amino acid differences have no effect on the sequence. To support Applicant's position and its new claims, Applicant submits scientific literature, attached hereto as Exhibit A. Exhibit A includes U.S. Patent No. 5,935,832; "Novel peptide ligands for integrin $\alpha_4\beta_1$ overexpressed in cancer cells," Mol Cancer Ther 2004; 3(10), October

2004; "Enzymatic characterization of *in vitro* of recombinant proprotein convertase PC4," Biochem J. (1999) 343, 29-37; "Identification of Calmodulin Isoform-specific Binding Peptides from a Phage-displayed Random 22-mer Peptide Library," The Journal of Biological Chemistry, Vol. 277, No.24, June 14, 2002, pp. 21630-21638; and "Peptide phosphorylation by calcium-dependent protein kinase from maize seedlings," Eur. J. Biochem, 267, 337-343 (2000).

The aforementioned and enclosed patent and scientific literature supports Applicant's traversal of the restriction/ election requirement whereby Applicant's deca-peptides are not independent and distinct from each other and that it is well known in the art to include generic or filler amino acids in sequences which can by filled by any amino acid. This concept is specifically denoted in Claim 1 of U.S. Patent No. 5,935,832, pages 1329, 1330, 1333 and 1334 of the "Novel peptide ligand" article, pages 337 and 342 of the "Peptide phosphorylation" article, pages 21630 and 21633 of the "Identification of Calmodulin" article and pages 29, 34, 36 and 37 of the "Enzymatic characterization" article. Therefore, because the use of generic amino acids is well known in the art and because claims 12, 14 and 15 are not independent and distinct, but rather equivalent, Applicant respectfully requests that Examiner withdraw its restriction/ election requirement and moreover, that he accept new claims 16 through 19 that read on the elected decapeptide and more particularly point out Applicant's invention. Further, because the use of generic amino acids is well known in the art, it is not new matter and the new claims are fully supported by the originally filed application.

Because the decapeptide in claim 13 is different from the sequences of claims 12, 14 and 15, Applicant cancels claim 13 without prejudice and reserves the right to file a divisional application directed to that decapeptide.

Conclusion

In view of the foregoing, Applicant respectfully requests that Examiner withdraw its restriction/ election requirement and further accept New Claims 16 through 19. With this amendment and response, no new matter has been added. Reexamination and reconsideration of this application, as amended, is requested. If any additional fees are required for this Petition, please deduct the required amounts from Deposit Account No. 500703 (Trojan Law Offices).

Respectfully Submitted,

TROJAN LAW OFFICES
BY

Dated: May 5, 2005

R. Joseph Trojan
Reg. No. 34,264

TROJAN LAW OFFICES
9250 Wilshire Blvd. Suite 325
Beverly Hills, CA 90212
Tel: (310) 777-8399
Fax: (310) 777-8348
Customer No. 23388

EXHIBIT A

US005935832A

United States Patent [19]

Nakane et al.

[11] Patent Number: 5,935,832

[45] Date of Patent: Aug. 10, 1999

[54] FARNESYL DIPHOSPHATE SYNTHASE

[75] Inventors: Hiroyuki Nakane; Chikara Ohto, both of Toyota; Shinichi Ohnuma, Sendai; Kazutake Hirooka, Sendai; Tokuzo Nishino, Sendai, all of Japan

[73] Assignee: Toyota Jidosha Kabushiki Kaisha, Toyota, Japan

[21] Appl. No.: 08/898,560

[22] Filed: Jul. 22, 1997

[30] Foreign Application Priority Data

Jul. 24, 1996 [JP] Japan 8-213211

[51] Int. Cl.[6] C12N 9/10; C12N 1/20; C12N 15/00; C07H 21/04

[52] U.S. Cl. 435/193; 435/252.3; 435/254.11; 435/325; 435/320.1; 435/410; 536/23.2; 935/22

[58] Field of Search 435/193, 320.1, 435/252.3, 254.11, 325, 410; 536/23.2; 935/22

[56] References Cited

FOREIGN PATENT DOCUMENTS

0 537 553 A2	4/1993	European Pat. Off. .
0 674 000 A2	9/1995	European Pat. Off. .
0 699 761 A2	3/1996	European Pat. Off. .
0 733 709 A2	9/1996	European Pat. Off. .
409065878	3/1997	Japan .

OTHER PUBLICATIONS

Math, et al., *Proc. Natl. Acad. Sci. USA*, vol. 89, Aug. 1992, pp. 6761–6764.

Ohnuma, et al., *J. Biol. Chem.*, 269:20, May 20 1994, pp. 14792–14797.

Koyama and Ogura, Mechanisms of Isoprenoid Chain Elongation Systems: Cloning and Analysis of Prenyltransferase Genes of *Bacillus stearothermophilus*, English language version of p. 174 of proceedings of the 36th Meeting for Natural Organic Compounds, Hiroshima–shi, Japan, Oct. 27–29, 1994, pp. 167–173.

Koyama, et al., *J. Biochem.*, 113:3, pp. 355–363, 1993.

Chen, et al., *J. Biochem.*, 268:15, pp. 11002–11007, 1993.

Chen, et al., *Arch. Biochem. Biophys.*, 314:2, pp. 399–404, Nov. 1 1994.

Chen, et al., *Protein Science*, vol. 3, pp. 600–607, 1994.

Anderson, et al., *J. Biol. Chem.*, vol. 264, pp. 19176–19184, Nov. 15 1989.

Jeong, et al., *J. DNA Sequencing & Mapping*, vol. 4, pp. 59–67, 1993.

Koike–Takeshita, et al., *J. Biol. Chem.*, 270:31, pp. 18396–18400, Aug. 4, 1995.

Fujisaki, et al., *J. Biochem*, 108:6, pp. 995–1000, 1990.

Ashby, et al., *J. Biol. Chem.*, 265:22, pp. 13157–13164, Aug. 5, 1990.

Teruya, et al., *Mol. Cell. Biol.*, 10:5, pp. 2315–2326, May 1990.

Sheares, et al. *Biochemistry*, 28:20, pp. 8129–8135, 1989.

Wilkin, et al., *J. Biol. Chem.*, 265:8, pp. 4607–4614, Mar. 15, 1990.

Armstrong, et a., *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 9975–9979, Dec. 1990.

Scolnik, et al., *Plant Physiol.*, vol. 104, pp. 1469–1470, 1994.

Aitken, et al., *Plant Physiol.*, vol. 108, pp. 837–838, 1995.

Badillo, et al., *Plant Mol. Biol.*, vol. 27, pp. 425–428, 1995.

Plant Physiol., 108:3, p. 1343, 1995.

Lang, et al., *J. Bacteriol.*, 177:8, pp. 2064–2073, Apr. 1995.

Armstrong, et al., *Mol. Gen. Genet.*, vol. 216, pp. 254–268, 1989.

EMBL release 45, Dec. 1995 (14 pages).

Carattoli, et al., *J. Biol. Chem.*, 266:9, pp. 5854–5859, Mar. 25, 1991.

Koyama, *Can. J. Chem.*, vol. 72, pp. 75–79, 1994.

Ohnuma, et al., *J. Biol. Chem.*, 271:17, pp. 10087–10095, Apr. 26, 1996.

Koyama, et al., *Biochemistry*, 33:42, pp. 12644–12648, 1994.

Ohnuma, et al., *J. Biol. chem.*, 271:31, pp. 18831–18837, Aug. 2, 1996.

Tarshis, et al., *Biochemistry*,33:36, pp. 10871–10877, 1994.

Koyama, et al., *Biochemistry*, 35:29, pp. 9533–9538, 1996.

Primary Examiner—Robert A. Wax
Assistant Examiner—Einar Stole
Attorney, Agent, or Firm—Kenyon & Kenyon

[57] ABSTRACT

A mutant prenyl diphosphate synthase capable of synthesizing prenyl diphosphates, shorter than those synthesized by the original enzyme, by modifying the amino acid sequence in and upstream of the aspartic acid-rich domain DDXX (XX)D (X denotes any amino acid, and XX in the parentheses may not be present) present in region II of the prenyl diphosphate synthase.

16 Claims, 3 Drawing Sheets

Fig. 1

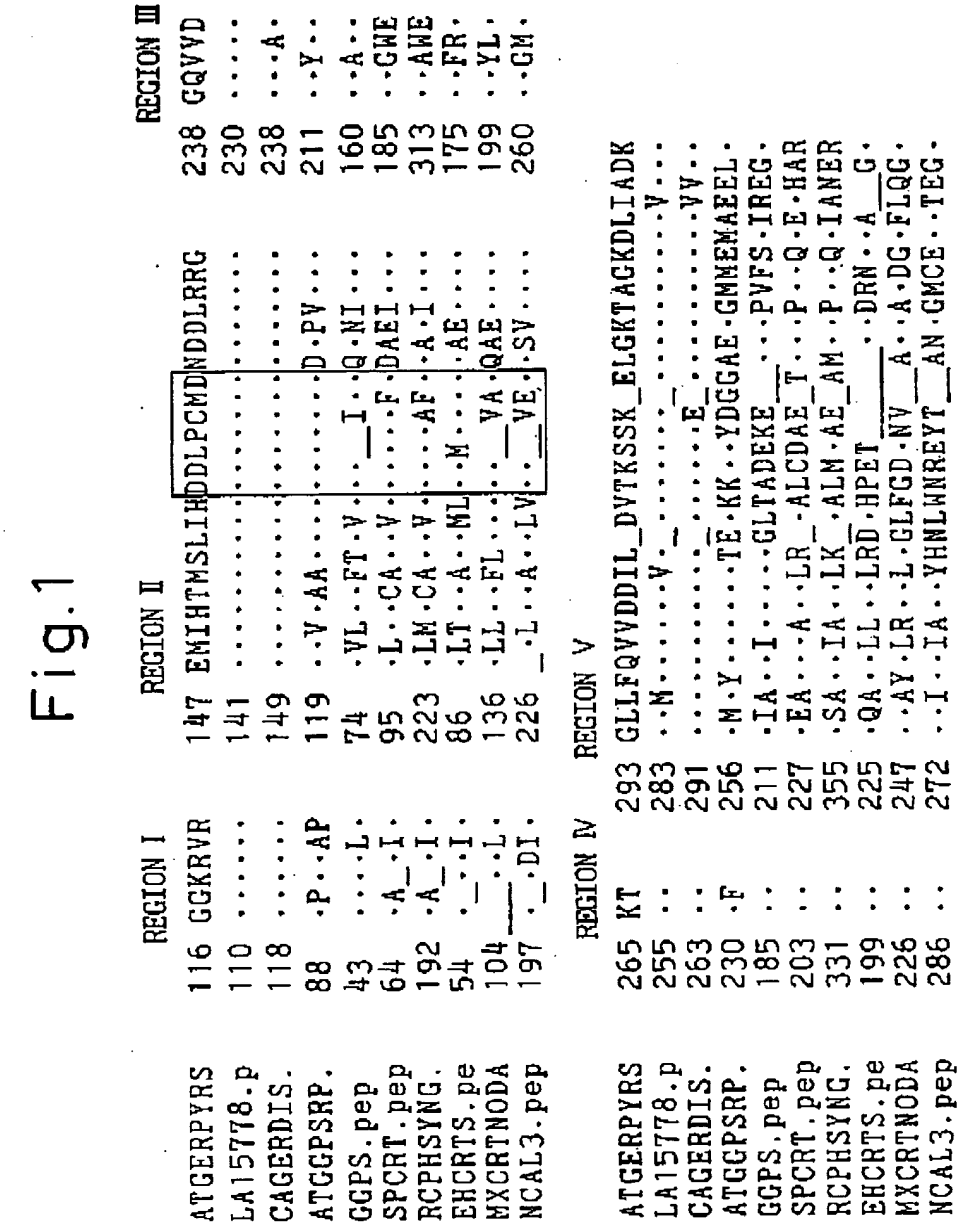

REGION I
REGION II
REGION III
REGION IV
REGION V
ATGERPYRS 116 GGKRVR 147 EMIHTMSLIHDDLPCMDNDDLRRG 238 GQVVD
LA15778.p 110 ..... 141 ..... 230 .....
CAGERDIS. 118 ..... 149 ..... 238 ---A-
ATGGPSRP. 88 -P--AP 119 --V-AA-----D-PV--- 211 --Y--
GGPS.pep 43 ---L- 74 -VL--FT-V---__I---Q-NI--- 160 --A--
SPCRT.pep 64 -A_-I- 95 -L--CA--V-----F-DAEI--- 185 --GWE
RCPHSYNG. 192 -A_-I- 223 -LM-CA--V-----AF--A-I--- 313 --AWE
EHCRTS.pe 54 -_--I- 86 -LT--A--ML---M-----AE---- 175 --FR-
MXCRTNODA 104 ___-L- 136 -LL--FL-----__VA-QAE---- 199 --YL-
NCAL3.pep 197 -_-DI- 226 _-L--A--LV---__VE--SV---- 260 --GM-
ATGERPYRS 265 KT 293 GLLFQVVDDIL_DVTKSSK_ELGKTAGKDLIADK
LA15778.p 255 -- 283 --N-------V-_---------_----------V---
CAGERDIS. 263 -- 291 ----------_-------E_----------VV--
ATGGPSRP. 230 -F 256 -M-Y------TE-KK--YDGGAE-GMMEMAEEL-
GGPS.pep 185 -- 211 -IA--I-----GLTADEKE___---PVFS-IREG-
SPCRT.pep 203 -- 227 -EA---A--LR_-ALCDAE_T---P--Q-E-HAR
RCPHSYNG. 331 -- 355 -SA--IA--LK_-ALM-AE_AM--P--Q-IANER
EHCRTS.pe 199 -- 225 -QA--LL--LRD-HPET_____--DRN--A___G-
MXCRTNODA 226 -- 247 --AY-LR--L-GLFGD-NV___A--A-DG-FLQG-
NCAL3.pep 286 -- 272 --I--IA--YHNLWNREYT__AN-GMCE--TEG-

Fig. 2

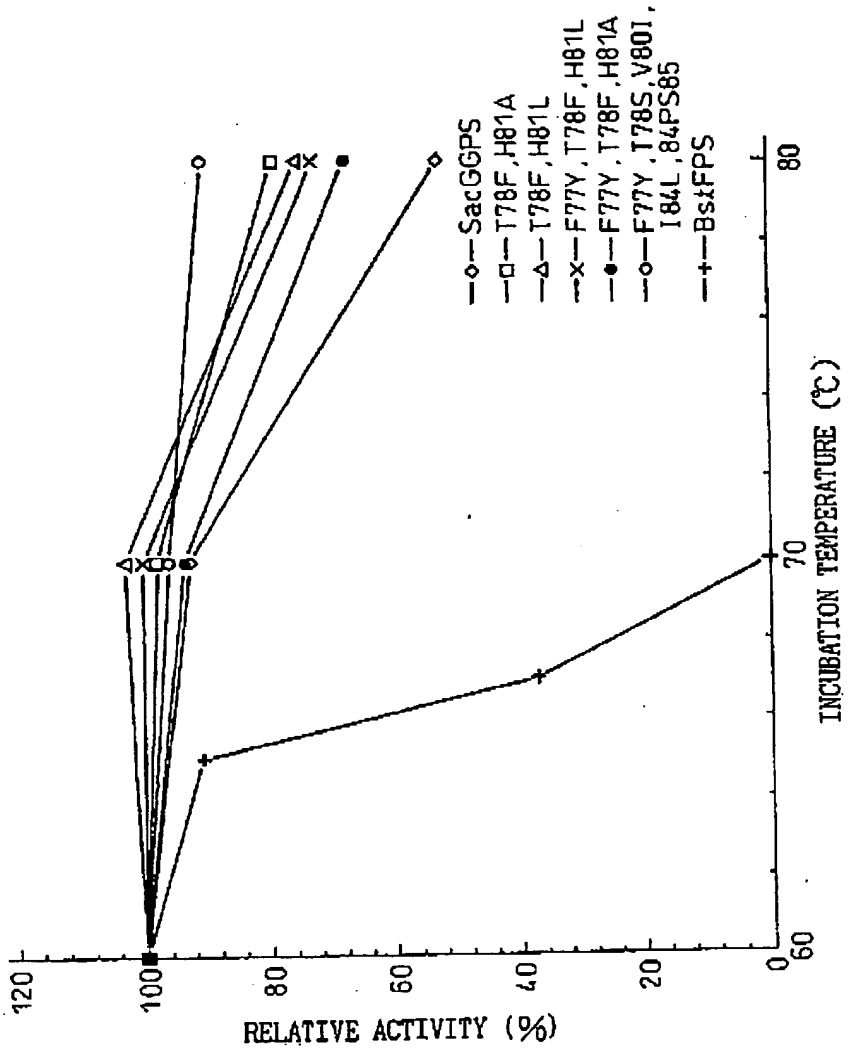
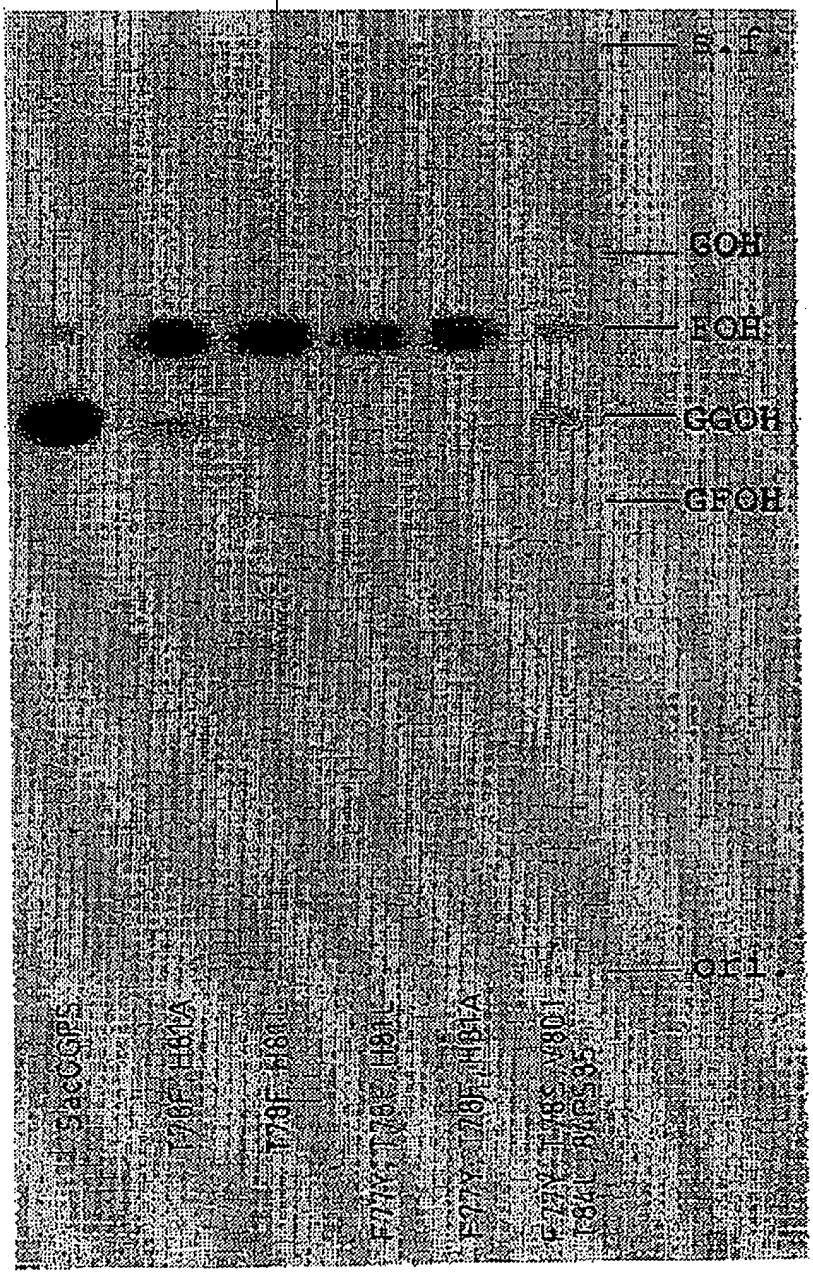

SacGGPS
T78F,H81A
T78F,H81L
F77Y,T78F,H81L
F77Y,T78F,H81A
F77Y,T78S,V80I,I84L,84PS85
BsiFPS
INCUBATION TEMPERATURE (℃)
60
70
80
RELATIVE ACTIVITY (%)
120
100
80
60
40
20
0

Fig. 3



FARNESYL DIPHOSPHATE SYNTHASE

BACKGROUND OF INVENTION

1. Field of Invention

The present invention relates to a novel mutant enzyme which synthesizes linear prenyl diphosphates that are precursors of compounds, important for organisms, such as steroids, ubiquinones, dolichols, carotenoids, prenylated proteins, animal hormones, plant hormones, and the like; a genetic system encoding said enzyme; and a method for producing and using said enzyme.

2. Related Art

Of the substances having important functions in organisms, many are biosynthesized using isoprene (2-methyl-1,3-butadiene) as a constituent units. These compounds are also called isoprenoids, terpenoids, or terpenes, and are classified depending on the number of carbon atoms into hemiterpenes (C5), monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), sesterterpenes (C25), triterpenes (C30), tetraterpenes (C40), and the like. The actual biosynthesis starts with the mevalonate pathway through which mevalonic acid-5-diphosphate is synthesized, followed by the synthesis of isopentenyl diphosphate (IPP) which is an active isoprene unit.

The identity of the isoprene unit that was proposed as a precursor was found to be isopentenyl diphosphate, the so-called active isoprene unit. Dimethylallyl diphosphate (DMAPP), an isomer of isopentenyl diphosphate, being used as a substrate in the synthesis of isopentenyl adenine which is known as a cytokinin, one of the plant hormones, it is also known to undergo a condensation reaction with isopentenyl diphosphate to synthesize chain-form active isoprenoids such as geranyl diphosphate (GPP), neryl diphosphate, farnesyl diphosphate (FPP), geranylgeranyl diphosphate (GGPP), geranylfarnesyl diphosphate (GFPP), hexaprenyl diphosphate (HexPP), heptaprenyl diphosphate (HepPP), and the like.

There are Z type and E type condensation reactions. Geranyl diphosphate is a product of E type condensation and neryl diphosphate is of Z type condensation. Although, the all-E type is considered to be the active form in farnesyl diphosphate and geranylgeranyl diphosphate, the Z type condensation reaction leads to the synthesis of natural rubber, dolichols, bactoprenols (undecaprenols), and plants various polyprenols found in. They are believed to undergo the condensation reaction using the phosphate ester bond energy of the pyrophosphate and the carbon backbone present in the molecule and to produce pyrophosphate as the byproduct of the reaction.

Farnesyl diphosphate or geranylgeranyl diphosphate serve as a reaction substrate leading to the synthesis of prenylated proteins (from farnesyl diphosphate or geranylgeranyl diphosphate) represented by G proteins that are important in the mechanism of signal transducer in the cell; cell membrane lipids (from geranylgeranyl diphosphate) of archaea; squalene (from farnesyl diphosphate) which is a precursor of steroids; and phytoene (from geranylgeranyl diphosphate) which is a precursor of carotenoids. Prenyl diphosphates from hexaprenyl diphosphate and heptaprenyl diphosphate having six and seven isoprene units, respectively, to prenyl diphosphates having ten isoprene units serve as the precursors of the synthesis of ubiquinone and menaquinone (vitamin K2) that work in the electron transport system.

Furthermore, via the biosynthesis of these active-form isoprenoids, a vast number of kinds of compounds that are

vital to life have been synthesized. Just to mention a few, there are cytokinins that are plant hormones and isopentenyl adenosine-modified tRNA that use hemiterpenes as their precursor of synthesis, geraniols and that isomer nerol belonging to monoterpens are the main components of rose oil perfume and a camphor tree extract, camphor, which is an insecticide. Sesquihormones include juvenile hormones of insects, diterpenes include a plant hormone gibberellin, trail pheromones of insects, and retinols and retinals that function as the visual pigment precursors, binding components of the purple membrane proteins of highly halophilic archaea, and vitamin A.

Furthermore, using squalene, a triterpene, a wide variety of steroid compounds have been synthesized, including, for example, animal sex hormones, vitamin D, ecdysone which is an ecdysis hormone of insects, a plant hormone brassinolide, constitution of the plasma membrane etc. Various carotenoids of tetraterpenes that are precursors of various pigments of organisms and vitamin A are also important compounds derived from active isoprenoids. Compounds such as chlorophyll, pheophytin, tocopherol (vitamin E), and phylloquinone (vitamin K1) are also derived from tetraterpenes.

The active isoprenoid synthases that sequentially condense isopentenyl diphosphates with such allylic substrates as dimethylallyl diphosphate, geranyl diphosphate, farnesyl diphosphate, geranylgeranyl diphosphate, geranylfarnesyl diphosphate, etc. are called the prenyl diphosphate synthases, and are also called, based on the name of the compound having the maximum chain length of the major reaction products, for example farnesyl diphosphate synthase (FPP synthase), geranylgeranyl diphosphate (GGPP synthase), and the like. There are reports on purification, activity measurement, genetic cloning, and sequencing of the DNA encoding enzymes such as farnesyl diphosphate synthase, geranylgeranyl diphosphate synthase, hexaprenyl diphosphate synthase, heptaprenyl diphosphate synthase, octaprenyl diphosphate synthase, nonaprenyl diphosphate synthase (solanesyl diphosphate synthase), undecaprenyl diphosphate synthase, and the like from bacteria, archaea, fungi, plants, and animals.

These active isoprenoid synthases constituting the basis of chemical synthesis of a great variety of compounds that are important both in the industry and in the academic field of life sciences have had few practical uses in the industrial application due to their unstable nature and low specific activities. However, with the isolation of thermostable prenyl diphosphate synthases from thermophilic bacteria and archaea and the genes encoding these enzymes, their availability as the enzyme has increased.

With regard to farnesyl diphosphate synthase, a gene was isolated from *Bacillus stearothermophilus*, a medium thermophile, and an enzyme having a medium thermal stability was prepared using *Escherichia coli* as host cell [T. Koyama et al. (1993) J. Biochem., 113: 355–363; Japanese Unexamined Patent Publication No. 5(1993)-219961]. With regard to geranylgeranyl diphosphate synthase, a gene was isolated from high thermophiles such as *Sulfolobus acidocaldarius* and *Thermus thermophilus* [S. -i. Ohnuma et al., (1994) J. Biol. Chem., 269: 14792–14797; Japanese Unexamined Patent Publication No. 7(1995)-308193, and; Japanese Unexamined Patent Publication No. 7(1995)-294956], and enzymes having a high thermal stability were prepared.

Furthermore, with regard to the prenyl diphosphate synthase having the functions of both of the farnesyl diphosphate synthase and the geranylgeranyl diphosphate synthase,

the enzyme and the gene encoding it have been isolated from highly thermophile *Methanobacterium thermoautotrophicum* [A. Chen and D. Poulter (1993) J. Biol. Chem., 268: 11002–11007; A. Chen and D. Poulter (1994) ARCHIVES OF BIOCHEMSTRY AND BIOPHYTSICS 314], and the thermostable nature of the enzyme has been demonstrated.

However, in the synthesis of farnesyl diphosphate/geranylgeranyl diphosphate derived from *Methanobacterium thermoautotrophicum*, there are no reports on the data of thin layer chromatography analysis etc. that can specify the chain length of the reaction products in connection with the assay of the enzymatic activity; the chain length has been estimated by measuring geranyl diphosphate as the allylic substrate. Since geranyl diphosphate can also serve as a substrate of geranylgeranyl diphosphate synthase, it is unlikely that the measured activity includes that of the farnesyl diphosphate synthase alone.

Moreover, the presence of farnesyl diphosphate synthase has not been confirmed in archaea that are expected to have enzymes having higher thermo stability, higher salt-stability and lower-pH-stability.

As mentioned above, the use of the farnesyl diphosphate synthase derived from *Bacillus stearothermophilus* resolved part of the problem of the enzyme being unstable and difficult to handle. But, an enzyme having a higher thermal stability would be more stable and more amenable to industrial application.

Moreover, some prenyl diphosphate synthases having a longer chain length use farnesyl diphosphate as a substrate. When such a long-chain prenyl diphosphate synthase is used simultaneously with a farnesyl diphosphate synthase for the purpose of providing the substrate of the former enzyme, the latter enzyme must have stability which is equal to or higher than that of the long-chain prenyl diphosphate synthase. When industrial production of farnesyl diphosphate is contemplated, the enzyme must be immobilized or recovered for recycling. When it is regenerated, the enzyme itself to be more stable, must have higher thermo stability, higher salt stability, and higher stability in a wider range of pH.

It has been found out that of the two aspartic acid-rich domains that have been proposed based on the amino acid sequence of the prenyl diphosphate synthase, the amino acid residue located at the fifth position in the N-terminal direction from the conserved sequence I (DDXX(XX)D) (wherein X denotes any amino acid, and the two X's in the parentheses may not be present) of the aspartic acid-rich domain in the amino-terminal side is responsible for controlling the chain length of the reaction product. Hence, a method has been invented that controls the reaction product for the purpose of lengthening the chain length of the reaction product [Japanese patent application No. 8-191635 filed on Jul. 3, 1996 under the title of "A Mutant Prenyl Diphosphate Synthase"]. The enzyme produced using the method enables production of reaction products that have several chain lengths. However, methods have not been not known that induce mutation of geranylgeranyl diphosphate synthase to control the reaction products to be in the short chain-length side in order to produce farnesyl diphosphate.

SUMMARY OF INVENTION

It is an object of the invention to establish a process for producing farnesyl diphosphate synthases by modifying amino acid sequences of prenyl diphosphate enzymes. A new enzyme that is more stable or that has a high specific activity more adaptable to industrial application would make it possible to obtain immediately a mutant prenyl diphosphate synthase or the gene thereof that produces farnesyl diphosphate and that retains the property owned by the the prenyl diphosphate synthase prior to mutation.

From the information on the nucleotide sequence of the gene of the geranylgeranyl diphosphate synthase of the mutant *Sulfolobus acidocaldarius* (*S. acidocaldarius*), it was clarified that out of the two Aspartic acid-rich domains that have been proposed based on the analysis of the amino acid sequence of prenyl diphosphate synthases, the amino acid residues within the aspartic acid-rich domain conserved sequence I (DDXX(XX)D) at the amino terminal side or the five amino acid residues to the N-terminal side from the amino terminal of said conserved sequence I are involved in the control of chain length of the reaction products.

Thus, the present invention provides a mutant prenyl diphosphate synthase having a modified amino acid sequence, wherein

at least one amino acid residue selected from (a) the amino acid residues in between the amino acid residue located at the fifth position in the N-terminal direction from D of the N-terminal and the amino acid residue located at the first position in the N-terminal direction from D of said N-terminal of the aspartic acid-rich domain DDXX(XX)D (wherein X sequence denotes any amino acid, and the two X's in the parentheses may not be present) present in region II, and (b) the amino acid residue located at the position in the N-terminal direction from D of the C-terminal of said aspartic acid-rich domain has been substituted by another amino acid, and/or

additional amino acid(s) have been inserted in between the amino acid residue located at the first position in the N-terminal direction from D of the C-terminal and D of said C-terminal of said aspartic acid-rich domain.

The present invention provides a farnesyl diphosphate-producing mutant prenyl diphosphate synthase which retains the properties that were owned by the native prenyl diphosphate synthase.

The present invention also provides a DNA or an RNA encoding the above enzyme.

The present invention further provides a recombinant vector and more specifically an expression vector comprising the above DNA.

The present invention further provides a host transformed by the above vector.

The present invention further provides a process for producing prenyl diphosphates having not more than 15 carbons comprising the step wherein the above enzyme is brought into contact with a substrate selected from the group consisting of isopentenyl diphosphate, dimethylallyl diphosphate, and geranyl diphosphate.

The present invention further provides a process of production of a mutant enzyme according to any of claims 1 to 8, said method comprising the steps of culturing the above host and of harvesting the expression product from the culture.

BRIEF EXPLANATION OF THE DRAWINGS

FIG. 1 is a graph showing the regions (I) to (V) and the aspartic acid-rich domain I of various prenyl diphosphate synthases. In the figure, the sequence represents the amino acid sequence of geranylgeranyl diphosphate synthase, and ATGERPYRS is the one derived from *Arabidopsis thaliana*, LA15778.p from *Lupinas olbus*, CAGERDIS from *Capsicum annuum*, ATGGPSRP from *Arabidopsis thaliana*, GGPS-pep from *Sulfolobus acidocaldarius*, SPCRT.pep

from *Rhodobacter sphaeroides*, RCPHSYNG from *Rhodobacter capsulatus*, EHCRTS.pe from *Erwinia herbicola*, MXCRTNODA from *Myxococcus thaliana*, and NCAL3.pep from *Neurospora crassa*. The number indicated on the left of each amino acid sequence represents the site from the N-terminal side of each geranylgeranyl diphosphate synthase at the N-terminal of the amino acid sequence.

FIG. 2 is a graph showing the thermal stability of the mutant prenyl diphosphate synthase. The ordinate shows the relative activity to 100% at incubation at 60° C. The abscissa shows the incubation temperature. SacGGPS is the geranylgeranyl diphosphate synthase prior to mutation. The others represent the mutant type enzyme of each. BstFPS is the farnesyl diphosphate synthase derived from *Bacillus stearothermophillus*.

FIG. 3 shows a photograph of a development pattern of thin layer chromatography of the dephosphorylated reaction products of the mutant prenyl diphosphate synthase when geranyl diphosphate was used as the allylic substrate. In the figure, ori. represents the origin of development, and s.f. represents the solvent front.

GOH is geraniol, FOH is farnesol, GGOH is geranyl geraniol, and GFOH is geranylfarnesol, and these are produced from dephosphorylation of geranyl diphosphate, farnesyl phosphate, geranylgeranyl diphosphate, and geranylfarnesyl diphosphate, respectively. SacGGPS is the geranylgeranyl diphosphate synthase prior to mutation. The others are each mutant enzymes.

DETAILED DESCRIPTION

It has been proposed that there are five conserved regions in the amino acid sequence of a prenyl diphosphate synthase (one subunit in the case of a heterodimer) [A. Chem et al., Protein Science Vol. 3, pp. 600–607, 1994]. It is also known that of the five conserved regions, there is an aspartic acid-rich domain conserved sequence I [DDXX(XX)D] (wherein X denotes any amino acid, and the two X's in the parentheses may not be present) in region II. Although there is also an aspartic acid-rich domain indicated as "DDXXD" in region V, the aspartic acid-rich domain used to specify the modified region of the amino acid sequence of the present invention is the one present in region II, and this domain is termed as the aspartic acid-rich domain I as compared to the aspartic acid-rich domain II present in region V.

As the prenyl diphosphate synthases having the aspartic acid-rich domain as described above, there can be mentioned farnesyl diphosphate synthase, geranylgeranyl diphosphate synthase, hexaprenyl diphosphate synthase, heptaprenyl diphosphate synthase, octaprenyl diphosphate synthase, nonaprenyl diphosphate synthase, undecaprenyl diphosphate synthase, and the like. More specific examples include the farnesyl diphosphate synthase of *Bacillus stearothermophilus*, the farnesyl diphosphate synthase of *Escherichia coli*, the farnesyl diphosphate synthase of *Saccharomyces cerevisiae*, the farnesyl diphosphate synthase of the rat, the farnesyl diphosphate synthase of the human, the geranylgeranyl diphosphate synthase of *Neurospora crassa*, the hexaprenyl diphosphate synthase of *Saccharomyces cerevisiae*, and the like.

By way of example of some of these, regions I to V and the aspartic acid-rich domain I (in the box) in region II of the amino acid sequence of geranylgeranyl diphosphate synthases are shown in FIG. 1.

The present invention can be applied to any prenyl diphosphate synthase having the aspartic acid-rich domain I.

In accordance with the present invention, in the amino acid sequence of a prenyl diphosphate synthase, at least one amino acid residue selected from (a) the amino acid residues in between the amino acid residue located at the fifth position in the N-terminal direction from D of the N-terminal and the amino acid residue located at the first position in the N-terminal direction from D of said N-terminal of the aspartic acid-rich domain DDXX(XX)D (wherein X denotes any amino acid, and the two X's in the parentheses may not be present) present in region II, and (b) the amino acid residue located at the first position in the N-terminal direction from D of the C-terminal of said aspartic acid-rich domain has been substituted by another amino acid, and/or

an additional one or more amino acids have been inserted in between the amino acid residue located at the first position in the N-terminal side from D of the C-terminal and D of said C-terminal of said aspartic acid-rich domain.

The mutant prenyl diphosphate synthase of the present invention can synthesize a farnesyl diphosphate having a shorter chain length than the prenyl diphosphate synthesized by the native prenyl diphosphate synthase.

In accordance with the present invention, by way of example, the gene of the geranylgeranyl diphosphate synthase of a highly thermophilic archaea, *Sulfolobus acidocaldarius*, is used as the starting material. *Sulfolobus acidocaldarius* is available from ATCC as ATCC No. 33909. The method for cloning the gene has been described in detail in Japanese Unexamined Patent Publication No. 7-308193. It has also been disclosed with the accession No. D28748 in the gene information data base such as GenBank. By using the sequence it can be cloned in the conventional method known in the art. An example of the other cloning methods is illustrated in Example 1 herein and its nucleotide sequence is shown as SEQ ID No: 2.

More specifically, the mutant enzyme of the present invention is a mutant prenyl diphosphate synthase characterized in that at least one amino acid selected from phenylalanine in position 77, threonine at position 78, valine at position 80, histidine at position 81, and isoleucine at position 84 has been substituted by another amino acid, and/or amino acid(s) have been inserted in between isoleucine at position 84 and methionine at position 85 in the geranylgeranyl diphosphate synthase having the amino acid sequence as set forth in SEQ ID No: 1.

By way of example, there is provided the amino acid sequences wherein the amino acids have been substituted as shown below:

Mutant enzyme 1: Changes from threonine at position 78 to phenylalanine, and histidine at position 81 to alanine;

Mutant enzyme 2: Changes from threonine at position 78 to phenylalanine, and histidine at position 81 to leucine;

Mutant enzyme 3: Changes from phenylalanine at position 77 to tyrosine, threonine at position 78 to phenylalanine, and histidine at position 81 to leucine;

Mutant enzyme 4: Changes from phenylalanine at position 77 to tyrosine, threonine at position 78 to phenylalanine, and histidine at position 81 to alanine;

Mutant enzyme 5: Changes from phenylalanine at position 77 to tyrosine, threonine at position 78 to serine, valine at position 80 to isoleucine, and isoleucine at position 84 to leucine, and an insertion of proline and serine in between isoleucine at position 84 and methionine at position 85.

In accordance with the present invention, it is indicated that the mutant prenyl diphosphate synthase retains the characteristic properties that were owned by the native prenyl diphosphate synthase. By way of example, the above-

mentioned five mutant enzymes show thermo resistance almost equal to that owned by the native geranylgeranyl diphosphate synthase.

It is known that an enzyme can sometimes exhibit its original enzymatic activity even when it has been modified by addition, removal, and/or substitution of one or a few amino acids as compared to the original amino acid sequence. Therefore, the present invention is intended to encompass those enzymes that have been modified by addition, deletion, and/or substitution of one or a few, for example up to five, or up to 10, amino acids as compared to the amino acid sequence as set forth in SEQ ID No: 1 and that can perform its original function.

The present invention also provides the genes encoding various mutant enzymes mentioned above, the vectors containing those genes, specifically expression vectors, and the hosts transformed by said vectors. The gene (DNA) of the present invention can be readily obtained, for example, by introducing mutation into the DNA encoding the original amino acid sequence as set forth in SEQ ID No: 1 using a conventional method such as site-directed mutagenesis, PCR and the like.

Furthermore, once the amino acid sequence of the desired enzyme has been determined, an appropriate nucleotide sequence encoding it can be determined, and the DNA can be chemically synthesized in accordance with a conventional method of DNA synthesis.

The present invention further provides an expression vector comprising DNA such as the one mentioned above, the host transformed by said expression vector, and a method for producing the enzyme or peptide of the present invention using these hosts.

Expression vectors contain an origin of replication, expression regulatory sequences etc., but they may differ depending on hosts used. As the hosts, there are mentioned procaryotes, for example, bacteria such as *Escherichia coli*, organisms of genus Bacillus such as *Bacillus subtilis*, and eukaryotic microorganisms, for example, fungi, for example yeast, for example *Saccharomyces cerevisiae* of genus Saccharomyces and *Pichia pastoris* of genus Pichia, filamentous fungi, for example the genus Asperaillus such as *Asperaillus niger*, animal cells, for example the cultured cells of the silkworm, cultured cells of higher animals, for example CHO cells, and the like. Furthermore, plants may also be used as the host.

As set forth in Examples, in accordance with the present invention, by cultivating the host transformed by the DNA of the present invention, farnesyl diphosphates may be accumulated in the culture broth, which may be harvested to produce their farnesyl diphosphates. Furthermore, in accordance with the invention, farnesyl diphosphates may also be produced by contacting the mutant prenyl diphosphate synthase produced by the method of the invention to the substrate isopentenyl diphosphate and each allyl substrate such as dimethylallyl diphosphate and geranyl diphosphate.

When *Escherichia coli* is used as the host, it is known that the host has the regulatory functions at the stage of transcribing mRNA from DNA and of translating protein from mRNA. As the promoter sequence regulating mRNA synthesis, in addition to the naturally occurring sequences (for example, lac, trp, bla, lpp, P_L, P_R, ter, T3, T7, etc.), there are known their mutants (for example, lac UV5), and the sequences (such as tac, trc, etc.) in which a naturally occurring promoter is artificially fused, and they can be used for the present invention.

It is known that the distance between the sequence of the ribosome biding site (GGAGG and similar sequences thereof) and the initiation codon ATG is important as the sequence regulating the ability of synthesizing protein from mRNA. It is also well known that a terminator (for example, a vector containing trn PT, T_2 is commercially available from Pharmacia) that directs transcription termination at the 3'-end affects the efficiency of protein synthesis by a recombinant.

As the vectors that can be used for preparation of the recombinant vectors of the present invention, commercially available vectors are used as they are, or various vectors may be mentioned that are derived depending on the intended use. For example, there can be mentioned pBR322, pBR327, pKK223-3, pKK233-3, pTrc99, and the like having a replicon derived from pMB1; pUC18, pUC19, pUC118, pUC119, pTV118N, pTV119N, pBluescript, pISG298, pHSG396, and the like that have been altered to enhance copy numbers; and pACYC177, pACYC184, and the like that have a replicon derived from p15A; and, furthermore, plasmids derived from pSC101, ColE1, R1, F factor, and the like.

Furthermore, fusion protein-expressing vectors that enable easier purification such as pGEX-2T, pGEX-3X, pMal-c2 may be used. One example of the gene used as the starting material of the present invention has been described in Japanese Unexamined Patent Publication No. 7-308193.

Furthermore, in addition to plasmids, virus vectors such as λ phage or M13 phage, or transposon may be used for the transformation of genes. With regard to the transformation of the gene into microorganisms other than *Escherichia coli*, gene transformation into organisms of genus Bacillus by pUB110 (commercially available from Sigma) or pHY300PLK (commercially available from Takara Shuzo) is known. These vectors are described in "Molecular Cloning" (J. Sambrook, E. F. Fritsch, and T. Maniatis, Cold Spring Harbor Laboratory Press) and "Cloning Vector" (P. H. Pouwels, B. B. Enger, Valk, and W. J. Brammar, Elsevier), and catalogues of the manufacturers.

Integration of the DNA fragment encoding the prenyl diphosphate synthase and, where needed, the DNA fragment having the function of regulating expression of the gene of said enzyme into these vectors can be performed by a known method using an appropriate restriction enzyme and ligase. Specific examples of the plasmids thus constructed include, for example, pBs-SacGGPS.

As the microorganisms into which genes can be directly introduced using such recombinant vectors include *Escherichia coli* and microorganisms of the genus Bacillus. Such transformation can also be carried out using general method, for example the $CaCl_2$ method and the protoplast method as described in "Molecular Cloning" (J. Sambrook, E. F. Fritsch, and T. Maniatis, Cold Spring Harbor Laboratory Press) and "DNA Cloning" Vol. I to III (D. M. Clover ed., IRL PRESS).

In order to produce the mutant enzyme of the present invention, a host transformed as above is cultured, and then said culture is subjected to any method comprising salting out, precipitation with an organic solvent, gel chromatography, affinity chromatography, hydrophobic chromatography, ion exchange chromatography, and the like to recover and purify said enzyme.

The present invention also provides a process for producing farnesyl diphosphates using the enzyme of the present invention. According to this process, the enzyme of the present invention is reacted with a substrate in a medium, particularly an aqueous medium, and then, as desired, the prenyl diphosphate is harvested from the reaction medium.

As the enzyme, not only a purified enzyme but also a crude enzyme that may be semi-purified to various stages, or a mixture of the cultured broth of a microorganism may be used. Alternatively there may be used immobilized enzymes prepared according to the general method from said enzyme, said crude enzyme, or a product containing the enzyme.

As the substrate, there may be used dimethyl allyl diphosphates or geranyl diphosphates and isopentenyl diphosphates. As the reaction medium, water or an aqueous buffer solution, for example Tris buffer or phosphate buffer and the like, may be used.

By using the method of producing the mutant prenyl diphosphate synthase obtained by the present invention, the mutant prenyl diphosphate synthase derived from a archaea may be created that is more stable and thus easier to handle and that produces prenyl diphosphate. Furthermore, there is also expected a creation of the farnesyl diphosphate-producing mutant prenyl diphosphate synthase that has the property of the prenyl diphosphate synthase prior to mutation (for example, salt stability or stability in a wide range of pH) added thereto.

In the claims and the specification of the present invention, amino acid residues are expressed by the one-letter codes or three-letter codes as described hereinbelow:

A; Ala; alanine
C; Cys; cysteine
D; Asp; aspartic acid
E; Glu; glutamic acid
F; Phe; phenylalanine
G; Gly; glycine
H; His; histidine
I; Ile; isoleucine
K; Lys; lysine
L; Leu; leucine
M; Met; methionine
N; Asn; asparagine
P; Pro; proline
Q; Gln; glutamine
R; Arg; arginine
S; Ser; serine
T; Thr; threonine
V; Val; valine
W; Trp; tryptophan
Y; Tyr; tyrosine

Substitution of amino acid is expressed in the order of "the amino acid residue before substitution," "number of the amino acid residue," and "the amino acid residue after substitution," by the one-letter codes of amino acids. For example, the mutation in which a tyrosine residue at position 81 is replaced with a methionine residue is expressed as Y81M. Furthermore, the insertion of amino acid residues is expressed by "the number of the amino acid residue at the N-terminal side of the insertion site prior to insertion," "the amino acid residue that was inserted," and "the number of the amino acid residue at the C-terminal side of the insertion site prior to insertion." For example, the insertion of alanine in between the amino acid at position 84 and the amino acid at position 85 is expressed as 84A85.

EXAMPLES

The present invention is now explained with reference to specific examples, but they must not be construed to limit the invention in any way.

Example 1

Construction of a Plasmid Containing the Gene for Geranylgeranyl Diphosphate Synthase

The gene for the geranylgeranyl diphosphate synthase (hereinafter referred to as SacGGPS) derived from *Sulfolobus acidocaldarius* was subcloned at the HindIII site of the plasmid vector pBluescript II (KS+) commercially available from Toyobouki. The plasmid DNA was designated as pBs-SacGGPS. The SacGGPS gene is available from *Escherichia coli* DH5α (pGGPS1) that was internationally deposited on Jan. 31, 1994 with the National Institute of Bioscience and Human Technology Agency of Industrial Science and Technology, of Ibaiaki, Japan under the accession number of FERM BP-4982.

Also, the entire nucleotide sequence of the SacGGPS gene has been published in Japanese Unexamined Patent Publication No. 7-308193 Shin-ichi Ohnuma et al. (1994) The Journal of Biological Chemistry Vol. 269:14792–14797, or in the genetic information data bank such as GenBank under the accession number D28748. Since *Sulfolobus acidocaldarius* is also available from various depositories of microorganisms such as ATCC etc. (as ATCC No. 33909), the DNA of the gene region of SacGGPS can be obtained by the conventional gene cloning method.

Example 2

Synthesis of the Oligonucleotides for Introducing Mutation

For introducing mutation of the gene of geranylgeranyl diphosphate synthase, the following oligonucleotides were designed and synthesized:

Primer DNA (T78F, H81A):
5'-CATACTTTTTTCCTTGTGGCTGATGATATCATG GATC-3' (SEQ ID No: 3)
Primer DNA (T78F, H81L):
5'-CATACTTTTTTCCTTGTGCTTGATGATATCATG GATC-3' (SEQ ID No: 4)
Primer DNA (F77Y, T78F, H81L):
5'-CATACTTATTTCCTTGTGCTTGATGATATCAT GGATC-3' (SEQ ID No: 5)
Primer DNA (F77Y, T78F, H81A):
5'-CATACTTATTTCCTTGTGGCTGATGATATCAT GGATC-3' (SEQ ID No: 6)
Primer DNA (F77Y, T78S, V80I, I84L, 84PS85):
5'-GTTCTTCATACTTATTCGCTTATTCATGATAGT ATT-3' (SEQ ID No: 7), and 5'-ATTCATGATGATC TTCCATCGATGGATCAAGAT-3' (SEQ ID No: 8).

Introduction of the mutation (F77Y, T78S, V80I, I84L, 84PS85) was effected using two nucleotides. First, mutation was introduced as mentioned in Example 3 using the oligonucleotide

5'-GTTCTTCATACTTATTCGCTTATTCATGATAG TATT-3' (SEQ ID No: 7) and a transformant was prepared in accordance with Example 4, and furthermore mutation was introduced into the plasmid thus obtained using the oligonucleotide

5'-ATTCATGATGATCTTCCATCGATGGATCAAGAT-3' (SEQ ID No: 8).

These nucleotides have a mutation in the codon encoding at least one amino acid residue selected from phenylalanine at position 77, threonine at position 78, valine at position 80, histidine at position 81, and isoleucine at position 84 in SacGGPS. In addition to the introduction of the codon

encoding an amino acid that has been inserted in between isoleucine at position 84 and methionine at position 85, they are designed to newly introduce the cleavage site of the restriction enzyme BspHI (5'TGATGA3'), the cleavage site of the restriction enzyme EcoRV (5'GATATC3'), or the cleavage site of the restriction enzyme ClaI (5'ATCGAT3'). In the introduction of the cleavage site of BspHI, the amino acid sequence encoded by the SacGGPS gene does not change due to degeneracy of codons, or it is a site for an introduction of mutation. This is used to detect the substitution-mutated plasmid by means of agarose gel electrophoresis after digestion with an appropriate restriction enzyme, since the introduction of mutation by substitution into the SacGGPS gene simultaneously produces new cleavage sites of restriction enzymes.

These primer DNA's were subjected to phosphorylation treatment at 37° C. for 30 minutes in the reaction medium shown below followed by denaturation at 70° C. for 10 minutes:

10 pmol/μl primer DNA 2 μl

10xkination buffer 1 μl

10 mM ATP 1 μl

H2O 5 μl

T4 polynucleotide kinase 1 μl

wherein the 10xkination buffer is 1000 mM Tris-Cl (pH 8.0), 100 mM $MgCl_2$, and 70 mM DTT.

Example 3

The Introduction of Substitution-Mutation of the SacGGPSS Gene

Using each primer DNA constructed in Example 2, substitution-mutation was introduced into the plasmid prepared in Example 1 in accordance with the Kunkel method. Mutan-K kit commercially available from Takara Shuzo was used to perform the Kunkel method. The experimental procedure was as described in the kit insert. The substitution-mutation of the plasmid need not be conducted by the Kunkel method. For example, an identical result can be obtained by a method using the polymerase chain reaction (PCR).

Using *Escherichia coli* CJ236 in the Mutan-K kit as the host cell, a single strand DNA was obtained in which a thymine base in plasmid pBS-SacGGPS was replaced with a deoxyuracil base.

The single strand DNA thus obtained was used as the template in the reaction in which a primer DNA for synthesizing a complementary strand was treated in the following reaction solution at 65° C. for 15 minutes and then annealed by allowing to stand at 37° C. for 15 minutes:

Single strand DNA 0.6 pmol

Annealing buffer solution 1 μl

Primer DNA solution (Example 2) 1 μl

H2O make to a final volume of 10 μl

in which the annealing buffer solution is 200 mM Tris-Cl (pH 8.0), 100 mM $MgCl_2$, 500 MM NaCl, and 10 mM DTT.

Furthermore, 25 μl of the elongation buffer solution, 60 units of *Escherichia coli* DNA ligase, and 1 unit of T4 DNA polymerase were added to synthesize the complementary strands at 25° C. for 2 hours. The elongation buffer solution is 50 mM Tris-Cl (pH 8.0), 60 mM ammonium acetate, 5 mM $MgCl_2$, 5 mM DTT, 1 mM NAD, and 0.5 mM dNTP.

After the reaction is over, 3 μl of 0.2 M EDTA (pH 8.0) was added thereto and was subjected to treatment at 65° C. for 5 minutes to stop the reaction.

Example 4

Construction of a Recombinant Having a Gene in Which Substitution-Mutation Has Been Introduced into the SacGGPS Gene

The DNA solution constructed in accordance with Example 3 was used to transform *Escherichia coli* XL1-Blue by the $CaCl_2$ method. An alternative method such as electroporation gives a similar result. A host cell other than *Escherichia coli* XL1-Blue, for example JM109 and the like also gave a similar result.

The transformant obtained by the $CaCl_2$ method was plated onto the agar plate containing ampicillin, a selectable marker of transformants, and was incubated overnight at 37° C.

Of the transformants obtained as above, the substitution-mutated pBs-SacGGPS plasmid that has a cleavage site of BspHI, EcoRV or ClaI was selected. The nucleotide sequence in the neighborhood of the codon corresponding to the amino acid residue that undergoes mutation of the SacGGPS gene of the selected substitution-mutated pBs-SacGGPS plasmid was determined by the dideoxy method. As a result, the pBs-SacGGPS plasmid containing the following five mutated SacGGPS genes was obtained. The nucleotide sequences encoding the amino acid sequences from the amino acid at position 77 to the amino acid at position 85 is shown below:

Mutation Nucleotide sequence

T77F, H81A: 5'-TTTTTCCTTGTGGCTGATGATATCATG-3' (SEQ ID No: 9)

T78P, H81L: 5'-TTTTTCCTTGTGCTTGATGATATCATG-3' (SEQ ID No: 10)

F77Y, T78F, H81L: 5'-TATTTCCTTGTGCTTGATGATATCATG-31 (SEQ ID No: 11)

F77Y, T78F, H81A: 5'-TATTTCCTTGTGGCTGATGATATCATG-3' (SEQ ID No: 12)

F77Y, T78S, V80I, I84L, 84PS85: 5'-TATTCGCTTATTCATGATGATCTTCCATCGATG-3' (SEQ ID No: 13)

Wild type: 5'-TTTACGCTTGTGCATGATGATATTATG-3' (SEQ ID No: 14).

Example 5

Measurement of Activity of the Mutant Prenyl Diphosphate Synthase

Crude enzyme solutions were prepared as follows from 6 transformants comprising 5 mutant SacGGPS genes and one wild type SacGGPS gene obtained in Example 4.

The transformant cultured overnight in the 2×LB medium was centrifuged to harvest cells, and then the cells were suspended into a buffer for cell homogenization (50 mM calcium phosphate buffer solution (pH 5.8), 10 mM β-mercaptoethanol, 1 mM EDTA). This was homogenized by sonication and then centrifuged at 4° C. at 10,000 r.p.m. for 10 minutes. The supernatant obtained was treated at 55° C. for 12 hours to inactivate the activity of prenyl diphosphate synthase derived from *Escherichia coli*. This was further centrifuged under the same condition and the supernatant obtained was used as a crude enzyme extract. When thermo stability was investigated the enzyme extract was incubated at 60° C., 70° C., or 80° C. (60° C., 65° C., 67° C., or 70° C. for the enzymes derived from *Bacillus stearothermophilus*) for one hour prior to reaction. The reaction was conducted at 55° C. for 15 minutes in the following reaction solution:

[1-^{14}C]-isopentenyl diphosphate (1 Ci/mol) 25 nmol
Allylic diphosphate (geranyl diphosphate) 25 nmol
Potassium phosphate buffer (pH 5.8) 10 mM
$MgCl_2$ 5 mM
Enzyme solution 100 μg
H_2O to make 200 μl

After the reaction is over, 200 μl of saturated NaCl was added to the reaction solution and 1 ml of water-saturated butanol was added thereto, which was then agitated, centrifuged, and separated into two phases. To 800 μl of the butanol layer obtained was added 3 ml of a liquid scintillator and then the radioactivity was measured by the scintillation counter. The result is shown in FIG. 2.

The mutant prenyl diphosphate synthase has exhibited a thermo stability which is equal to that of the native geranylgeranyl diphosphate synthase, and is higher than that of the farnesyl diphosphate synthase derived from *Bacillus stearothermophilus*.

The solvent is evaporated from the remainder of the butanol layer by purging nitrogen gas thereinto while heating the layer in order to concentrate to a volume of about 0.5 ml. To the concentrate were added 2 ml of methanol and one ml of potato acid phosphatase solution (2 mg/ml potato acid phosphatase, 0.5 M sodium acetate (pH 4.7)) to effect the dephosphorylation reaction at 37° C. Subsequently the dephosphorylated reaction product was extracted with 3 ml of n-pentane.

This was concentrated by evaporating the solvent by purging nitrogen gas thereinto, which was then analyzed by TLC (reverse phase TLC plate: LKC18 (Whatman), development solvent: acetone/water=9/1). The developed dephosphorylated reaction product was analyzed by the Bio Image Analyzer BAS2000 (Fuji Photo Film) to determine the location of radioactivity. The result when geranyl diphosphate was used as the allylic substrate is shown in FIG. 3.

The reaction product of the mutant prenyl diphosphate synthase was shown to be a farnesyl diphosphate.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 14

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Sulfolobus acidocaldarius
(B) STRAIN: ATCC 33909

(ix) FEATURE:
(A) NAME/KEY: Asp-rich domain
(B) LOCATION: 82-86

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
 1               5                  10                  15

Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
            20                  25                  30

Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
        35                  40                  45

Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
    50                  55                  60

Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
65                  70                  75                  80

His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
                85                  90                  95

Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
            100                 105                 110

Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
        115                 120                 125

Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
    130                 135                 140

Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
```

-continued

```
145                 150                 155                 160

Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
                165                 170                 175

Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
            180                 185                 190

Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
        195                 200                 205

Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
    210                 215                 220

Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
225                 230                 235                 240

Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
                245                 250                 255

Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
            260                 265                 270

Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
        275                 280                 285

Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
    290                 295                 300

Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
305                 310                 315                 320

Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
                325                 330
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 993 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Sulfolobus acidocaldarius
(B) STRAIN: ATCC 33909

(ix) FEATURE:
(A) NAME/KEY: Asp-rich domain coding
(B) LOCATION: 246-258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
ATGAGTTACT TTGACAAGTA TTTTAATGAG ATTGTTAATT GTGTAAACGA CATTATTAAG     60

AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT    120

GCAGCTAACA GGTTAAGACC ATTAATCTTA AGTATATCAT CAGATTTATT CGGAGGACAG    180

AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTGATACTTT TACCCTTGTG    240

CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA    300

TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC    360

TTAACCCAGG CTCTTAGAGG TTTCCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC    420

AGTCGTTCAA TAATAATTAT ATCCGAAGGA CAGCCACTAG ATATGGAATT TGAGGACAGA    480

ATTGATATAA AGGAGCACGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC    540

TCGGCATGCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG    600

ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT    660

CTAACACCAG ACGAAAAGGA ACTTCGAAAG CGTGTTTTTA GTGATATTAG GGAGGGTAAA    720
```

-continued

AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACCAGAA CAAGATTGTC 780

CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGACCTC AGCAGATATA 840

ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT 900

ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTCGAAAGGC TTTAAAATAT 960

CTAGCTGAAT TTACGATAAG AAGGACAAAA TAA 993

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATACTTTTT TCCTTGTGGC TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CATACTTTTT TCCTTGTGCT TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CATAGTTATT TCCTTGTGCT TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CATACTTATT TCCTTGTGGC TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTTCTTCATA CTTATTCGCT TATTCATCAT AGTATT 36

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATTCATGATC ATCTTCCATC CATGGATCAA CAT 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTTTTCCTTG TGGCTGATGA TATCATG 27

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTTTCCTTG TGCTTGATGA TATCATG 27

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TATTTCCTTG TGCTTGATGA TATCATG 27

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TATTCCTTG TGGCTGATGA TATCATG 27

-continued

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TATTCGCTTA TTCATCATGA TCTTCCATCG ATG 33

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTTACGCTTC TGCATCATGA TATTATC 27

We claim:

1. A mutant prenyl diphosphate synthase having a modified amino acid sequence, wherein

said mutant prenyl diphosphate synthase comprises an aspartic acid-rich domain having the sequence, $D_1D_2X_1X_2(X_3X_4)D_3$, in region II of said mutant prenyl diphosphate synthase,

wherein each of D_1, D_2 and D_3 denote an aspartic acid residue; X_1, X_2, X_3 and X_4 are each independently any amino acid and X_3 and X_4 are each optionally independently present in the aspartic acid rich domain, and wherein

said mutant prenyl diphosphate synthase comprises (1) at least one amino acid substitution, said at least one amino acid substitution located at at least one amino acid position selected from (a) an amino acid between D_1 and the amino acid residue at the fifth position upstream of D_1 and (b) the amino acid residue located one amino acid position upstream of D_2; (2) at least one additional amino acid inserted between D_2 and the first amino acid upstream of D_3; or a combination of (2) and (3);

wherein said mutant prenyl diphosphate synthase synthesizes prenyl diphosphate which is shorter than prenyl diphosphate synthesized by a corresponding wild-type enzyme.

2. A mutant prenyl diphosphate synthase according to claim 1 wherein said mutant has the enzymatic activities and thermo stability of wild type prenyl diphosphate synthase.

3. A mutant enzyme according to claim 1 wherein the reaction product of the prenyl diphosphate synthase is farnesyl diphosphate.

4. A mutant enzyme according to claim 1 wherein the prenyl diphosphate synthase is of the homodimer-type.

5. A mutant enzyme according to claim 1 wherein the prenyl diphosphate synthase is derived from archaea.

6. A mutant enzyme according to claim 1 wherein the prenyl diphosphate synthase is derived from *Sulfolobus acidocaldarius*.

7. A mutant enzyme according to claim 1 wherein the prenyl diphosphate synthase is a thermostable enzyme.

8. A mutant prenyl diphosphate synthase according to claim 1, wherein at least one amino acid selected from phenylalanine at position 77, threonine at position 78, valine at position 80, histidine at position 81, and isoleucine at position 84 has been substituted by another amino acid, or one or more amino acids have been inserted in between isoleucine at position 84 and methionine at position 85 in the geranylgeranyl diphosphate synthase as set forth in SEQ ID No: 1.

9. A mutant prenyl diphosphate synthase according to claim 1 wherein at least one amino acid selected from phenylalanine at position 77, threonine at position 78, valine at position 80, histidine at position 81, and isoleucine at position 84 has been substituted by another amino acid, and/or two amino acids have been inserted in between isoleucine at position 84 and methionine at position 85 in the geranylgeranyl diphosphate synthase as set forth in SEQ ID NO: 1, wherein the phenyl alanine at position 77 has been replaced with tyrosine, the threonine at position 78 has been replaced with phenylalanine or serine, the valine at position 80 has been replaced with isoleucine, the histidine at position 81 has been replaced with leucine or alanine, or the isoleucine at position 84 has been replaced with leucine; or proline and serine have been inserted in between the isoleucine at position 84 and the methionine at position 85.

10. A mutant prenyl diphosphate synthase according to claim 1, wherein the mutant prenyl diphosphate synthase is derived from a native geranylgeranyl diphosphate synthase of an organism selected from the group consisting of *Arabidopsis thaliana, Lupinas albus, Capsicum annuun, Sulfolobus acidocaldarius, Rhodobactor sphaeroides, Rhodobactor capsulatus, Erwinia herbicola, Myxococcus thaliana* and *Neurospora crassa*.

11. A DNA encoding an enzyme according to claim 1.

12. An RNA transcribed from a DNA according to claim 11.

13. A recombinant vector comprising a DNA according to claim 11.

14. A host organism transformed with a recombinant vector according to claim 13.

15. A process for producing a mutant enzyme according to claim 1, said method comprising the steps of culturing a host transformed with an expression vector comprising a DNA coding for the mutant enzyme and of harvesting the expression product from the culture.

16. A process for producing a prenyl diphosphate having not more than 15 carbons comprising the step of bringing an enzyme according to claim 1 or any of claims 2 to 10 or an enzyme produced by the method according to claim 15 into contact with a substrate selected from the group consisting of isopentenyl diphosphate, dimethylallyl diphosphate, and geranyl diphosphate.

* * * * *

Novel peptide ligands for integrin $\alpha_4\beta_1$ overexpressed in cancer cells

Masahito Mikawa, Henry Wang, Linlang Guo, Ruiwu Liu, Jan Marik, Yoshikazu Takada, Kit Lam, and Derick Lau

University of California at Davis Cancer Center, Sacramento, California

Abstract

Using the "one-bead one-peptide" combinatorial technology, a library of random cyclic octapeptides and nonapeptides, consisting of natural and unnatural amino acids, was synthesized on polystyrene beads. This library was used to screen for peptides that promoted attachment and proliferation of bronchioloalveolar carcinoma cells (H1650), employing a "cell growth on bead" assay. Consensus peptide sequences of cNleDXXXXc and cXNleDXXXXc (where Nle is norleucine) were identified. With alanine scanning and site-directed deletion, a typical ligand consisted of a motif of -NleDI/V/Nle- with two flanking cysteines. These peptide ligands were specific for promoting cell attachment of the H1650 cells and the cells of lymphoid cancers (Jurkat and Raji) but not other selected human cell lines of lung cancer and fibroblast. In an antibody blocking assay, integrin $\alpha_4\beta_1$, which was overexpressed in H1650, Jurkat, and Raji, was identified as a putative receptor for these peptide ligands. Using Chinese hamster ovary cells transfected with either wild-type or mutant integrin α_4, a critical binding site for these peptides was localized to the glycine residue at position 190 of integrin α_4. [Mol Cancer Ther 2004;3(10):1329–34]

Introduction

Lung cancer is one of the most common cancers and is the leading cause of cancer death in the United States (1). The outlook for patients with lung cancer is poor, because they commonly present with advanced disease and their median survival is ~10 months even with chemotherapy (2, 3). Molecularly targeted therapy has become a reality in the treatment of several cancers (4). Rituximab, a monoclonal antibody targeting the cell surface antigen CD20, is being used to treat B-cell lymphoma (5). Trastuzumab, a monoclonal antibody targeting the cell surface receptor of HER-2/*neu* (erbB-2), has been used for treatment of breast cancer (6). More recently, an antibody against the epidermal growth factor receptor, cetuximab, has been approved for treatment of refractory colon cancer (7). We reason that unique surface receptor(s) also exist for lung cancer cells and that peptide ligands specific for these cell surface receptors could be identified. Our hope is to identify these peptide ligands and to develop them as diagnostic and therapeutic agents for this devastating disease (8).

Recently, using the technology of combinatorial chemistry and a "cell growth on bead" assay, we have identified a novel cyclic peptide, D-cysteine-L-asparagine-L-glycine-L-arginine-L-glycine-L-glutamate-L-glutamine-D-cysteine (cNGRGEQc), specific for promoting attachment and growth of cells of non-small cell lung cancer on polyethylene glycol–linked polystyrene beads (9). Furthermore, we have identified integrin $\alpha_3\beta_1$ as a putative cell surface receptor for this peptide ligand. While screening for peptide ligands specific for lung cancer cells, we have also identified a group of novel cyclic peptides specific for integrin α_4 overexpressed in a cell line of bronchioloalveolar carcinoma, a histologic subtype of lung cancer. We are reporting the identification and characterization of these novel peptide ligands for promoting adhesion of cancer cells, which overexpressed integrin α_4.

Materials and Methods

Cell Cultures

Human cell lines including H1650 (bronchioloalveolar carcinoma), A549 (lung adenocarcinoma), Calu-1 (squamous lung carcinoma), DMS-114 (small cell lung carcinoma), Jurkat (T-cell leukemia), Raji (Burkitt lymphoma), and Hs68 (newborn foreskin fibroblast) were obtained from American Tissue Culture Collection (Manassas, VA). The Chinese hamster ovary (CHO) lines have been described by Irie et al. and were obtained from Dr. Y. Takada (University of California at Davis, Sacramento, CA). The lung cancer and fibroblast cell lines were maintained in DMEM, the CHO cells in DMEM in sodium pyruvate (110 mg/mL) and G418 (100 μg/mL), and the Jurkat and Raji cells in RPMI 1640. All the culture media were supplemented with fetal bovine serum (10%), penicillin (100 units/mL), and streptomycin (100 μg/mL; Life Technologies, Grand Island, NY).

Random Peptide Bead Library

An "one-bead one-peptide" combinatorial library, containing TentaGel S-NH_2 polystyrene beads (Rapp Polymere GmbH, Tubingen, Germany) grafted with amino-polyethylene glycol and conjugated with random peptides of a general structure of NH_2-c(X)$_n$c-COOH-, was synthesized using Fmoc chemistry and a split-synthesis approach as

Received 7/16/03; revised 8/11/04; accepted 8/27/04.

Grant support: Pamela Pond Scholarship and Joan's Legacy Foundation grant (http://www.joanslegacy.org).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Derick Lau, Division of Hematology/Oncology, University of California at Davis Cancer Center, 4501 X Street, Sacramento, CA 95817. Phone: 916-734-3772; Fax: 916-734-7946. E-mail: derick.lau@ucdmc.ucdavis.edu

described previously by Lam et al. (10, 11). For this random library, *n* represents six to seven amino acids in length; c denotes D-cysteine at the NH_2 and COOH termini providing intramolecular cyclization by disulfide bonding; and X denotes 1 of 17 natural L-amino acids, excluding L-cysteine, L-arginine, and L-lysine, and 7 unnatural amino acids of L-hydroxyproline (Hyp), L-α-aminoisobutyric acid (Aib), L-norleucine (Nle), L-phenylglycine, L-norvaline, L-cyclohexylglycine, and L-4-benzoylphenylalanine. (Note: L-cysteine was excluded to avoid intramolecular cyclization, and the positive charged amino acids, L-arginine and L-lysine, were excluded to minimize nonspecific cell attachment.) Therefore, for a random combinatorial library of $c(X)_{6-7}c$ synthesized from 24 amino acids, there were, theoretically, $24^7 = 4.8 \times 10^9$ permutations of peptide sequences.

"Cell Growth on Bead" Assay

A "cell growth on bead" assay was used as described previously for screening the random library for peptides essential for attachment and growth of the H1650 cells on the TentaGel beads (9). Cultured cells were harvested with trypsin/EDTA, washed, and resuspended as single cells in culture medium. Typically, ~150,000 beads with random peptides were mixed with ~1 million suspended cells in 15 mL of medium distributed in six 3-cm culture dishes. The dishes were incubated in a tissue culture incubator with 5% CO_2 at 37°C for 48 to 72 hours. Beads covered with a monolayer of cells were removed and attached cells were stripped off with 8 mol/L guanidine hydrochloride. The amino acid sequence of each isolated peptide bead was determined with an automated Procise 494 protein sequencer (Applied Biosystems, Foster City, CA) using a method as described previously by Liu and Lam (12). Consensus peptide sequences identified were synthesized on beads in large quantity and cell attachment to these beads was confirmed with H1650. The structures of these peptides were modified to yield a model ligand peptide, D-cysteine-L-Nle-L-aspartate-L-Nle-L-threonine-L-Hyp-D-arginine-D-cysteine (cNleDNleTHyprc or pM2), with high capacity and specificity for promoting attachment of the H1650 cells and with resistance to proteolysis. The pM2 beads were subsequently used to test for cell type specificity and structure-activity relationship.

Alanine Scanning and Site-Directed Deletion Analysis

To determine the structure-activity requirements of the consensus peptide ligand, the strategies of alanine scanning (13) and site-directed deletion were employed. The model ligand peptide, pM2, was synthesized with Fmoc chemistry (11) on TentaGel beads with each amino acid (except the cysteine) sequentially substituted with L-alanine. In addition, truncated pM2 peptides, with amino acids deleted sequentially from the COOH terminus, were synthesized on beads. Percentage of the peptide beads with cell attachment was evaluated with the H1650 cells using the "cell growth on bead" assay.

Antibody Blocking Assay

To determine whether a particular integrin was a putative receptor for pM2, a panel of antibodies against α and β integrins, including α_1 to α_6 and α_v and β_1 to β_5 (Chemicon, Temecula, CA), was tested to block cell attachment to the pM2 beads. Suspended H1650 cells were incubated with each antibody at 37°C with agitation for 30 minutes. Peptide beads were then added to the culture, which was incubated at 37°C for 72 hours in a tissue culture incubator. The number of beads with cell attachment was counted under a microscope and relative percentage of beads with cell attachment was calculated as follows: $B / A \times 100$, where *A* is percentage of beads with cell attachment in medium without antibody and *B* is percentage of beads with cell attachment in medium with antibody.

Inhibition Assay with Soluble Peptides

To determine if pM2 was biologically active from the beads, free pM2 was synthesized and tested for competitive inhibition of attachment of H1650 to the model pM2 beads. The peptides of pM2 and its analogue of cNleANleTHyprc (pM2-A), the latter containing A instead of D at position 3 to serve as a negative control, were synthesized on rink amide-MBHA resins (GL Biochem Ltd., Shanghai, China) with Fmoc chemistry as described above. The peptides were subsequently cleaved from the resins and the free peptides were purified as described previously (14). The free pM2 and pM2-A were readily soluble in an aqueous solution. Inhibition of the soluble peptides on attachment of H1650 to the pM2 and pM2-A beads was done as described for the antibody blocking assay.

Flow Cytometry for Detection of α_4 and β_1 Integrins

Flow cytometry was used to compare the levels of expression of α_4 and β_1 integrins among the selected cell lines as described previously (15). Cells were incubated with a mouse anti-human α_4 (Serotec, Raleigh, NC) or β_1 antibody (Chemicon), 2 μg/mL, for 30 minutes on ice. After washing, the cells were incubated with a FITC-conjugated goat anti-mouse IgG antibody (Jackson ImmunoResearch Laboratories, West Grove, PA) for 30 minutes. Cells were analyzed for level of fluorescence using a Coulter EPICS XL flow cytometer (Beckman Coulter, Miami, FL).

Adhesion to pM2 Beads of CHO Cells with Wild-type or Mutant Integrin α_4

To determine the potential binding site of α_4 integrin for pM2 in initiating cell attachment, the model of CHO cells with transfected wild-type (CHO-α_4) or mutant human integrin α_4 was employed. This model was established by Irie et al. by transfecting CHO, a cell line devoid of endogenous integrin α_4 expression, with cDNA of human wild-type integrin α_4 or integrin α_4 substituted with alanine sequentially (alanine scanning mutagenesis) from residues 108 to 268. These residues are believed to contain the putative binding sites for vascular cell adhesion molecule-1 and fibronectin (16, 17). The CHO cell lines harboring the mutant integrin α_4 at residues 176 to 203 were tested in this study. The levels of expression of integrin α_4 in the CHO cells and CHO-α_4 cells were verified with flow cytometry as described above. Cell adhesion to the pM2 beads of these CHO cells was studied using the "cell growth on bead" assay.

Results

Peptide Ligands for H1650

Using the "cell growth on bead" assay to screen a "one-bead one-peptide" random library of cyclic octapeptides and nonapeptides for attachment of H1650 cells, 72 beads with cell attachment were isolated from screening ~4.7 million of peptide beads. The peptide sequences of these beads were determined and 12 beads with the most common consensus sequences of amino acids were identified: cNleDIDLLc, cNleDITPLc, cGNleDINleHypAibc, cFNleDIDAHc, cNleDINleEFGc, cFNleDIYCSc, cFNleDINeGTc, cNleDNleYEAibc, cNleDNleTHypGc, cNleDNleHPHypMc, cNleDVVDTSc, and cNleDVDNleHGc. These consensus peptide beads were synthesized in large quantities and their capacity of promoting cell attachment and growth was confirmed with the H1650 cells. These sequences contained a motif of -NleDX-, where D is aspartic acid and X is isoleucine, Nle, or valine. Attachment of one to three cells to the peptide beads was observable within 6 hours (Fig. 1). By transferring beads with attached cells to a fresh culture plate, cell proliferation on the beads was observed in 24 hours and confluency of cells on beads was commonly observed in 96 to 120 hours (Fig. 2).

To obtain a peptide homologous with the peptide ligands identified in above but with enhanced capacity of attracting cell attachment and with resistance to proteolytic digestion, the peptide sequence of D-cysteine-L-Nle-L-aspartate-L-Nle-L-threonine-L-Hyp-glycine-D-cysteine (cNleDNleTHypGc or pM1) was modified with the glycine replaced with D-arginine to yield the peptide of cNleDNleTHypγc or pM2. The modified peptide, pM2, was resistant to proteolytic digestion in human plasma. In addition, pM2 showed higher capacity of promoting cell attachment than pM1. Cell type specificity of attachment to pM2 beads was tested with a panel of human cell lines. Cell attachment was consistently observed with H1650, Jurkat (T-cell leukemia), and Raji (Burkitt lymphoma) but not with A549 (lung adenocarcinoma), Calu-1 (squamous lung carcinoma),

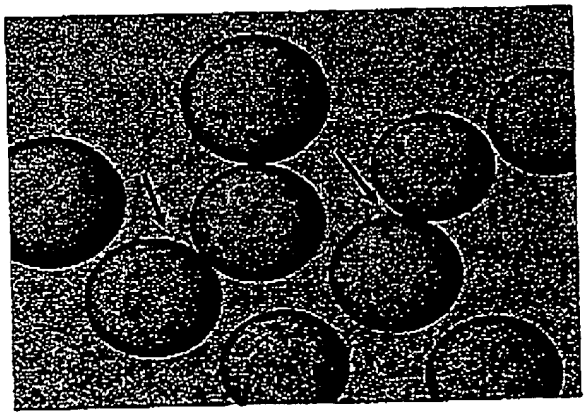


Figure 1. Attachment of H1650 cells to TentaGel beads conjugated with the pM2 peptide, cNleDNleTHypγc, after 6 hours of incubation. *Arrow*, attachment of single cell to a bead.

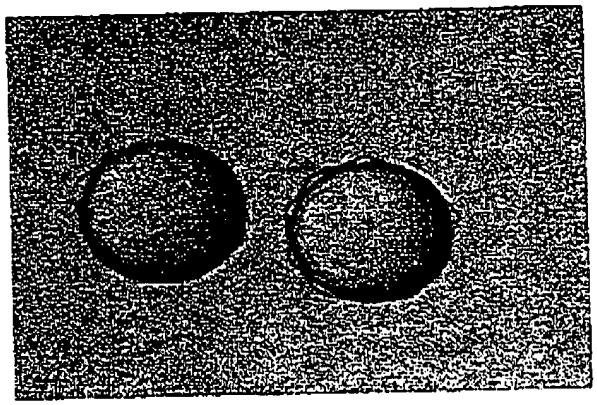


Figure 2. Peptide beads isolated from Fig. 1 and incubated for 96 hours. Note that the beads are covered with cells, indicating cell proliferation on beads.

DMS-114 (small cell lung carcinoma), and Hs68 (fibroblast; Table 1). It was hypothesized that these findings were due to overexpressions of α_4 and β_1 integrins in H1650 cells as compared with that of other lung cancers. To prove this hypothesis, flow cytometry was used to compare the levels of expression of α_4 and β_1 integrins among the selected cell lines. Cell attachment to the pM2 beads correlated with the level of expression of integrin α_4 (Table 1). Integrin α_4 was expressed at a relatively high level with the H1650, Jurkat, and Raji cells but at a low level with A549, Calu-1, and DMS-114 cells. Relative high levels of integrin β_1 were detected in all these cell lines (data not shown).

Structure-Activity Relationship Studies

By alanine scanning, the amino acids of L-Nle at position 2 and L-aspartate (D) at position 3 of pM2 were absolutely required for promoting cell attachment of H1650 to beads (Table 2). Substitution of amino acids at positions 4 to 6 with L-alanine essentially did not affect cell

Table 1. Cell type–specific attachment to pM2 beads and relative level of integrin α_4 expression of human cell lines

Cell Lines	Tumor Types	Cell Attachment	Integrin α_4*
H1650	Lung bronchioloalveolar carcinoma	Yes	71.8
A549	Lung adenocarcinoma	No	9.8
Calu-1	Squamous cell lung carcinoma	No	8.4
DMS-114	Small cell lung carcinoma	No	8.9
Hs68	Foreskin fibroblast	No	7.0
Jurkat	T-cell leukemia	Yes	135.9
Raji	Burkitt lymphoma	Yes	48.0

*Mean fluorescent intensity as detected with flow cytometry. Mean of two experiments.

Table 2. Cell attachment of H1650 to TentaGel beads conjugated with the peptide pM2 and its analogues modified with alanine scanning (mean ± SD, $n = 3$)

Peptide Sequences	Cell Attachment (%)
D-Cys-Nle-Asp-Nle-Thr-Hyp-D-Arg-D-Cys (pM2)	100
D-Cys-Ala-Asp-Nle-Thr-Hyp-D-Arg-D-Cys	10.6 ± 3.0
D-Cys-Nle-Ala-Nle-Thr-Hyp-D-Arg-D-Cys	0
D-Cys-Nle-Asp-Ala-Thr-Hyp-D-Arg-D-Cys	100
D-Cys-Nle-Asp-Nle-Ala-Hyp-D-Arg-D-Cys	100
D-Cys-Nle-Asp-Nle-Thr-Ala-D-Arg-D-Cys	100
D-Cys-Nle-Asp-Nle-Thr-Hyp-Ala-D-Cys	100

attachment. By site-directed deletion, the extent of attachment of the H1650 cells to the native and modified pM2 beads is shown in Table 3. The extent of cell attachment decreased progressively as amino acids were deleted sequentially from the COOH terminus. Furthermore, deletion of D-cysteine from the NH_2 terminus almost abolished cell attachment to most beads. The peptide of cNleDNlec maintained ~50% of biological activity. However, peptide of cNleDc lost almost all of its activity. Thus, the structure-activity requirements for pM2 are summarized as follows: (*a*) The motif for cell attachment is NleDNle. (*b*) The cyclic structure provided by the two D-cysteines is essential for biological activity. (*c*) The minimal length of the peptide for promoting cell attachment is three amino acids with two flanking cysteines.

Identification of Putative Receptor for Peptide Ligand

We hypothesized that integrin was the receptor for pM2 for promoting cell attachment. A panel of antibodies to α and β integrin subunits, including α_1 to α_6, α_v, and β_1 to β_5, were tested to block attachment of H1650 cells to the pM2 beads. At a concentration of 1 μg/mL, antibody to α_4 totally blocked cell attachment, and antibody to β_1 only partially blocked cell attachment to the peptide beads. The other antibodies did not affect cell attachment at all. The combination of antibodies to α_4 and β_1 subunits was additive in blocking cell attachment. The antibody to α_4 blocked cell attachment in a dose-dependent manner as shown in Fig. 3. The concentration of α_4 antibody that inhibited attachment of cells to 50% of the peptide beads (IC_{50}) was 2.2 ng/mL. The IC_{50} for the antibody to β_1 could not be determined because this antibody blocked cell attachment only partially even at a concentration up to 1 μg/mL. These data indicated that integrin $\alpha_4\beta_1$ was likely a receptor for the pM2 peptide ligand.

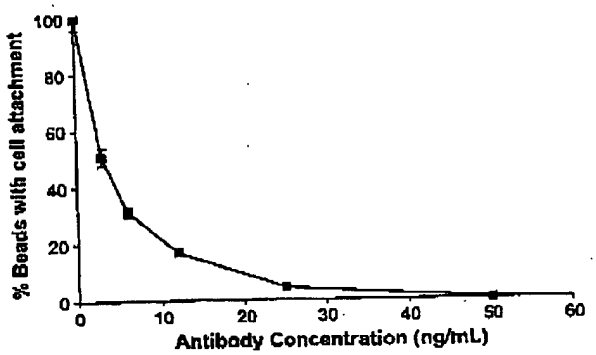


Figure 3. Dose-response curve of inhibition of cell attachment to pM2 beads by an antibody to integrin α_4 (mean ± SD, $n = 4$).

Biological Activity of Soluble pM2 Peptide

The biological activity of soluble pM2 and pM2-A peptides was tested by competitive inhibition of attachment of H1650 cells to the pM2 beads. The pM2-A peptide at concentration as high as 1 mmol/L did not inhibit cell attachment at all. The soluble pM2 inhibited cell attachment to the pM2 beads in a concentration-dependent manner as shown in Fig. 4. The concentration of soluble pM2 that inhibited attachment of H1650 cells to 50% of the pM2 beads (IC_{50}) was 10 μmol/L.

Identification of Binding Site of Integrin α_4 for pM2

The profile of cell attachment to the pM2 beads for CHO, CHO-α_4, and the mutant CHO-α_4 created by alanine scanning mutagenesis is shown in Fig. 5. Essentially no

Table 3. Cell attachment of H1650 to TentaGel beads conjugated with the peptide pM2 and its analogues modified by sequential deletion of amino acids from the COOH terminus (mean ± SD, $n = 3$)

Peptide Sequences	Cell Attachment (%)
D-Cys-Nle-Asp-Nle-Thr-Hyp-D-Arg-D-Cys (pM2)	100 ± 0
D-Cys-Nle-Asp-Nle-Thr-Hyp-D-Cys	87 ± 6
D-Cys-Nle-Asp-Nle-Thr-D-Cys	64 ± 1
D-Cys-Nle-Asp-Nle-D-Cys	47 ± 5
D-Cys-Nle-Asp-D-Cys	5 ± 4
Nle-Asp-Nle-Thr-Hyp-D-Arg-D-Cys	5 ± 2
Nle-Asp-Nle	1 ± 1

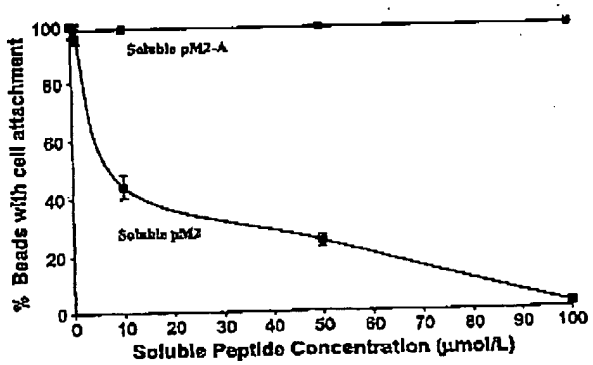


Figure 4. Dose-response curve of inhibition of cell attachment to pM2 beads by the soluble pM2 and pM2-A peptides (mean ± SD, $n = 4$).

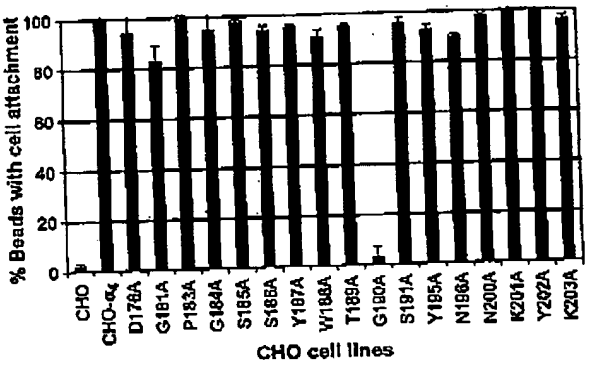


Figure 5. Profile of cell attachment to pM2 beads of parental CHO, CHO-α_4, and CHO-α_4 mutants created by alanine scanning mutagenesis at residues 176 to 203 of integrin α_4 (mean ± SD, $n = 4$).

attachment to pM2 beads was observed with the parental CHO cells. On the other hand, cell attachment to all the pM2 beads was observed with the CHO-α_4. For the CHO-α_4 mutants, cell attachment was observed on 80% to 100% of the pM2 beads. The only exception is the G190A mutant, in which the glycine was replaced with alanine at residue 190 of integrin α_4. These findings indicated that this glycine residue was critical for binding of pM2 to integrin α_4 for promoting cell attachment.

Discussion

Peptide ligands for various cell surface receptors have been identified using phage libraries. The best known peptide identified by this technique is the tripeptide of arginine-glycine-aspartate, which is a ligand for $\alpha_v\beta_3$ and $\alpha_5\beta_1$ integrins (18, 19). The sequence of arginine-glycine-aspartate has been identified as an essential site of several natural extracellular matrix, including vitronectin, laminin, fibrinogen, and fibronectin, for binding to their receptors (20–23). Whereas the technique of phage display provides the advantage of identification of ligands for cellular receptors *in vivo*, the "one-bead one-peptide" method, on the other hand, provides the ease of *in vitro* study and allows the use of random peptides containing unnatural amino acids (8). In this study, we have identified peptide ligands containing D- and synthetic unnatural amino acids. These peptides are resistant to proteolytic enzymes, which very likely would enhance biological stability *in vivo*. The inclusion of these unnatural amino acids also exponentially increases the number of permutations of random peptides. In our study, the library of random cyclic octapeptides and nonapeptides contained $24^6 + 24^7 = (1.91 \times 10^8 + 4.59 \times 10^9) \simeq 4.8 \times 10^9$ possible permutations of peptide sequences using 24 different amino acids. Of 4.7 million peptide beads screened in this study, which represented only a small fraction of possible permutations of the library, we isolated 72 beads with cell attachment. On sequencing, 12 peptides with consensus motifs of -NleDI-, -NleDNle-, and -NleDV- were identified. Using the techniques of alanine walk and sequential deletion, the motif of -NleDI/V/Nle- was confirmed. The glycine residue of position 190 of integrin α_4 has been identified as a potential binding site for the pM2 model peptide. It is likely that additional motifs could have been identified if more random peptide beads had been screened.

Cell type–specific peptide ligands for lung cancer cells have been identified using the "one-bead one-peptide" combinatorial library and the "cell growth on bead" assay established in our laboratory. Previously, we reported that a cyclic peptide, cNGRGEQc, was identified for promoting adhesion of non-small cell lung cancer cells overexpressing integrin $\alpha_3\beta_1$ (9). In this study, we have additionally identified a series of cyclic octapeptides and nonapeptides with a general structure of cNleDXXXXc or cXNleDXXXXc in a bronchioloalveolar carcinoma cell line, H1650, overexpressing $\alpha_4\beta_1$. Because there is a lack of availability of appropriate bronchioloalveolar carcinoma cell lines, we were unable to verify whether these peptides were specific for bronchioloalveolar carcinoma or $\alpha_4\beta_1$ expression was unique in this cell type of lung cancer. In an immunohistochemical study by Koukoulis et al. (24) of seven cases of bronchioloalveolar carcinoma lung cancer, no expression of α_4 was observed, but strong expression of integrin β_1 was detected. Additional studies with a larger sample size are necessary to confirm these observations.

Although it is uncertain if the pM2 peptide is specific for bronchioloalveolar carcinoma cells in general, it is obvious that this peptide is specific for cancer cells overexpressing integrin α_4. This notion is supported by the fact that the pM2 peptide specifically promoted attachment of integrin α_4 expressing cells of H1650 (bronchioloalveolar carcinoma), Jurkat (T-cell leukemia), and Raji (Burkitt lymphoma). Recently, a series of peptides containing the motifs of LDI/V/F has also been identified for $\alpha_4\beta_1$ integrin of Jurkat cells on screening "one-bead one-peptide" libraries (25). The pM2 analogues identified in the current study seem to have pharmacokinetic advantages because they contain D- and unnatural amino acids and are more resistant to proteolytic digestion in human plasma. Thus, these peptides may be more ideal for development into diagnostic and therapeutic agents.

In this and previous studies, we have identified cyclic peptides specific for integrins $\alpha_3\beta_1$ and $\alpha_4\beta_1$. These peptides are cell type specific depending on the expression of a particular type of integrin. In the clinical setting, bronchioloalveolar carcinoma, adenocarcinoma, and squamous carcinoma have been classified under the general category of non-small cell lung cancer and they, traditionally, have been treated in a similar manner. However, it is becoming clear that not all non-small cell lung cancers are alike and that molecular signature targets may provide a more accurate guidance for classification and specific therapy of lung cancers. Cell type–specific expression of integrins such as α_3 or α_4 seems to be viable diagnostic and therapeutic targets. In a recent clinical trial, a response rate of only 10% was

observed for non-small cell lung cancer treated with gefitinib (Iressa), an inhibitor of the tyrosine kinase domain of epidermal growth factor receptor (26). It has been found recently that the response to gefitinib was observed mostly in lung cancer harboring mutations of in-frame deletion or substitutions clustered around the ATP binding pocket of the tyrosine kinase binding domain of epidermal growth factor receptor (27). Therefore, gene profiling has been advocated as a more accurate means than the classic histologic observation in classifying and guiding treatment for lung cancer (28, 29).

Preclinical studies have shown the promise and feasibility of cancer therapy with antibodies and radiolabeled arginine-glycine-aspartate peptide targeting integrins (30, 31). In addition, liposome-encapsulated doxorubicin loaded with the Fab' fragment of an antibody to β_1 integrin has been used to target human lung cancer xenografts (32, 33). To study if the peptides with a general structure of cNleDXXXXc indeed are of diagnostic and therapeutic value, we have shown that soluble pM2, but not the control peptide pM2-A, is capable of inhibiting attachment of H1650 to the pM2 beads. Studies are ongoing to further define the biological activity of soluble pM2 *in vitro* and *in vivo*.

References

1. Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. Cancer statistics, 2003. CA Cancer J Clin 2003;53:5 – 26.

2. Kelly K, Crowley J, Bunn P Jr, et al. Randomized phase III trial of paclitaxel plus carboplatin versus vinorelbine plus cisplatin in the treatment of patients with advanced non-small-cell lung cancer: a Southwest Oncology Group trial. J Clin Oncol 2001;19:3210 – 8.

3. . Schiller JH, Harrington D, Belani CP, et al. Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. N Engl J Med 2002;346:92 – 8.

4. Goldman JM, Melo JV. Targeting the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med 2001;344:1084 – 6.

5. Davis TA, Grillo-Lopez AJ, White CA, et al. Rituximab anti-CD20 monoclonal antibody therapy in non-Hodgkin's lymphoma: safety and efficacy of re-treatment. J Clin Oncol 2000;18:3135 – 43.

6. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001;344:783 – 92.

7. Ellis L, Hoff PM. Targeting the epidermal growth factor receptor: an important incremental step in the battle against colorectal cancer. J Clin Oncol 2004;22:1177 – 9.

8. Aina OH, Sroka TC, Chen ML, Lam KS. Therapeutic cancer targeting peptides. Biopolymers 2002;66:184 – 99.

9. Lou DH, Guo L, Liu R, Song A, Shao C, Lam KS. Identifying peptide ligands for cell surface receptors using cell-growth-on-bead assay and one-bead one-compound combinatorial library. Biotechnol Lett 2002;24: 497 – 500.

10. Lam KS, Salmon SE, Hers EM, Hruby VJ, Kazmierski WM, Knapp RJ. A new type of synthetic peptide library for identifying ligand-binding activity. Nature 1991;354:82 – 4.

11. Lam KS, Lebl M, Krchnak V. The "one-bead-one-compound" combinatorial library method. Chem Rev 1997;97:411 – 48.

12. Liu R, Lam KS. Automatic Edman microsequencing of peptides containing multiple unnatural amino acids. Anal Biochem 2001;295:9 – 16.

13. Cunningham BC, Wells JA. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science 1989; 244:1081 – 5.

14. Lam KS, Lehman AL, Song A, et al. Synthesis and screening of "one-bead one-compound" combinatorial peptide libraries. Methods Enzymol 2003;369:298 – 322.

15. Cardarelli PM, Yamagata S, Taguchi I, Gorcsan F, Chiang SL, Lobl T. The collagen receptor $\alpha_2\beta_1$, from MG-63 and HT1080 cells, interacts with a cyclic RGD peptide. J Biol Chem 1992;267:23159 – 64.

16. Iric A, Kamata T, Puzon-McLaughlin W, Takada Y. Critical amino acid residues for ligand binding are clustered in a predicted β-turn of the third N-terminal repeat in the integrin α_4 and α_5 subunits. EMBO J 1995;14:5550 – 6.

17. Irie A, Kamata T, Takada Y. Multiple loop structures critical for ligand binding of the integrin α_4 subunit in the upper face of the β-propeller mode 1. Proc Natl Acad Sci 1997;94:7198 – 203.

18. Koivunen E, Wang B, Ruoslahti E. Phage libraries displaying cyclic peptides with different ring sizes: ligand specificities of the RGD-directed integrins. Biotechnology 1995;13:265 – 70.

19. Koivunen E, Gay DA, Ruoslahti E. Selection of peptides binding to the $\alpha_5\beta_1$ integrin from phage display library. J Biol Chem 1993;268:20205 – 10.

20. Ruoslahti E, Hayman EG, Pierschbacher MD. Extracellular matrices and cell adhesion. Arteriosclerosis 1985;5:581 – 94.

21. Ruoslahti E. RGD and other recognition sequences for integrins. Annu Rev Cell Dev Biol 1996;12:697 – 715.

22. Pierschbacher MD, Ruoslahti E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Nature 1984;309:30 – 3.

23. Hayman EG, Pierschbacher MD, Ruoslahti E. Detachment of cells from culture substrate by soluble fibronectin peptides. J Cell Biol 1985; 100:1948 – 54.

24. Koukoulis GK, Warren WH, Virtanen I, Gould VE. Immunolocalization of integrins in the normal lung and in pulmonary carcinomas. Hum Pathol 1997;28:1018 – 25.

25. Park S, Renil M, Vikstrom B, et al. The use of one-bead one-compound combinatorial library method to identify peptide ligands for $\alpha_4\beta_1$ integrin receptor in non-Hodgkin's lymphoma. Lett Peptide Sci 2002;8:171 – 8.

26. Kris MG, Natale RB, Herbst RS, et al. Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. JAMA 2003; 290:2149 – 58.

27. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129 – 39.

28. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M. Diversity of gene expression in adenocarcinoma of the lung. Proc Natl Acad Sci 2001;98:13784 – 9.

29. Gordon GJ, Jensen RV, Hsiao LL, et al. Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. Cancer Res 2002;62:4963 – 7.

30. Janssen ML, Oyen WJ, Dijkgraaf I, et al. Tumor targeting with radiolabeled α(v)β(3) integrin binding peptides in a nude mouse model. Cancer Res 2002;62:6146 – 51.

31. van Hagen PM, Breeman WA, Bernard HF, et al. Evaluation of a radiolabeled cyclic DTPA-RGD analogue for tumor imaging and radionuclide therapy. Int J Cancer 2000;90:186 – 98.

32. Abra RM, Bankert RB, Chen F, et al. The next generation of liposome delivery systems: recent experience with tumor-targeted, sterically-stabilized immunoliposomes and active-loading gradients. J Liposome Res 2002;12:1 – 3.

33. Sugano M, Egilmez NK, Yokota SJ, et al. Antibody targeting of doxorubicin-loaded liposomes suppresses the growth and metastatic spread of established human lung tumor xenografts in severe combined immunodeficient mice. Cancer Res 2000;60:6942 – 9.

Eur. J. Biochem. 267, 337–343 (2000) © FEBS 2000

Peptide phosphorylation by calcium-dependent protein kinase from maize seedlings

Mart Loog[1,2], Reet Toomik[2], Katrin Sak[1], Grazyna Muszynska[3], Jaak Järv[1] and Pia Ek[2]

[1]*Institute of Chemical Physics, Tartu University, Estonia;* [2]*Department of Medical Biochemistry and Microbiology, Uppsala University, Sweden;* [3]*Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland*

Ca^{2+}-dependent protein kinase (CDPK-1) was purified from maize seedlings, and its substrate specificity studied using a set of synthetic peptides derived from the phosphorylatable sequence RVLSRLHS[15]VRER of maize sucrose synthase 2. The decapeptide LARLHSVRER was found to be efficiently phosphorylated as a minimal substrate. The same set of peptides were found to be phosphorylated by mammalian protein kinase Cβ (PKC), but showed low reactivity with protein kinase A (PKA). Proceeding from the sequence LARLHSVRER, a series of cellulose-membrane-attached peptides of systematically modified structure was synthesised. These peptides had hydrophobic (Ala, Leu) and ionic (Arg, Glu) amino acids substituted in each position. The phosphorylation of these substrates by CDPK-1 was measured and the substrate specificity of the maize protein kinase characterised by the consensus sequence motif $A/L_{-5}X_{-4}R_{-3}X_{-2}X_{-1}SX_{+1}R_{+2}Z_{+3}R_{+4}$, where X denotes a position with no strict amino acid requirements and Z a position strictly not tolerating arginine compared with the other three varied amino acids. This motif had a characteristic sequence element RZR at positions +2 to +4 and closely resembled the primary structure of the sucrose synthase phosphorylation site. The sequence surrounding the phosphorylatable serine in this consensus motif was similar to the analogous sequence K/RXXS/TXK/R proposed for mammalian PKC, but different from the consensus motif RRXS/TX for PKA.

Keywords: Ca^{2+}-dependent protein kinase (CDPK); maize seedlings; protein kinase C (PKC); SPOTs peptides.

Ca^{2+}-dependent protein kinases (CDPKs) of plant origin form a distinct group of protein kinases [1], and the rather scarce studies have pointed to their possible involvement in regulation of metabolism and gene expression [2–4]. Plant CDPK molecules consist of a single polypeptide with a typical protein kinase catalytic core, a linker region containing the autoinhibitory sequence and an adjoining calmodulin-like domain with four calcium-binding EF-hand motifs [5]. The two latter structural fragments are responsible for regulation of CDPK activity by cytosolic Ca^{2+} ions, which act similarly to mammalian cells as second messengers in plants. However, in contrast with mammalian Ca^{2+}-dependent protein kinases [protein kinase C (PKC) and the calcium/calmodulin-dependent protein kinase], there is still not much known about the physiological substrates of plant CDPKs and the relevant signalling pathways.

Recently it has been shown that one of the key enzymes of sucrose metabolism in maize, sucrose synthase type 2, can be phosphorylated at Ser15 *in vivo* [6,7]. This reaction seems to have regulatory significance, as the catalytic properties of the phosphorylated enzyme were affected [6].

The phosphorylatable serine residue of sucrose synthase 2 is surrounded by the peptide sequence RVLSRLHS[15]VRER, with several positively charged amino acids before and after the Ser15 residue. This means that the substrate specificity of the relevant plant protein kinase(s) should follow similar principles to that of a large group of mammalian protein kinases, showing selectivity for positively charged amino acids [8]. The most extensively investigated members of this group are PKC and protein kinase A (PKA).

In the present work, the substrate specificity of a protein kinase purified from maize seedlings was investigated using synthetic peptides derived from the maize sucrose synthase 2 phosphorylation site sequence, which were phosphorylated by this kinase in a Ca^{2+}-dependent manner. The influence of peptide length on phosphorylation effectiveness was studied, and the sequence LARLHSVRER was found to be the minimal peptide substrate for this enzyme. Using this sequence, a series of structurally varied analogues were synthesized as cellulose-membrane-attached (SPOTs™ membranes) peptides. In these peptides, all the positions were substituted with hydrophobic (Leu, Ala) and charged (Arg, Glu) amino acids. By systematically varying the peptide structure, we were able to obtain information on the specificity pattern of maize CDPK, and this was compared with that of the mammalian protein kinases PKC and PKA.

MATERIALS AND METHODS

Materials

[γ-^{32}P]ATP was purchased from Amersham International. Histone H1 (type IIIS), pepstatin A and phenylmethanesulfonyl fluoride were products of Sigma. The SPOTs™ kit was purchased from Cambridge Research Biochemicals (Northwich, Cheshire, UK). Fmoc amino acids, supplied as active esters, were included in the kit or purchased from Millipore. DEAE-cellulose (DE-52) and 3MM ion-exchange paper were

Correspondence to J. Järv, Institute of Chemical Physics, Tartu University, 2 Jakobi St., 51014 Tartu, Estonia. Fax: + 372 7375 247, Tel.: + 372 7375 246, E-mail: jj@chem.ut.ee

Abbreviations: CDPK, Ca^{2+}-dependent protein kinase; PKA, protein kinase A; PKC, protein kinase C.

Enzyme: Ca^{2+}-dependent protein kinase (EC 2.7.1.37).

(Received 4 October 1999; accepted 9 November 1999)

from Whatman. The SMART system, phenyl-Sepharose and the Mono Q column were from Pharmacia. Polyacrylamide/bis (29 : 1), low-molecular-mass reference proteins and Coomassie Brilliant Blue R-250 were obtained from Bio-Rad. All other chemicals were of the highest purity commercially available. The peptides used were all synthesised in the Department of Medical Biochemistry and Microbiology at the University of Uppsala.

Partial purification of CDPK from maize seedlings

Corn seeds (*Zea mays* var. Mona) were soaked for 12 h in distilled water. They were grown in the dark for 72 h at 26 °C. The apical parts of the seedlings, which reached a length of about 2.5 cm, were harvested, immediately frozen in liquid nitrogen and stored at −70 °C. For the purification, 10 g of frozen seedlings were ground in a mortar to a fine powder in the presence of a small amount of liquid nitrogen and extracted with 30 mL homogenisation buffer (20 mM Tris/HCl, pH 7.5, 250 mM sucrose, 2 mM EDTA, 2 mM EGTA, 0.7 $\mu g \cdot mL^{-1}$ pepstatin A, 0.5 mM dithiothreitol, and 0.26 mM phenylmethanesulfonyl fluoride). All steps in this purification were performed at 4 °C. Tris base (1 M) was rapidly added dropwise to adjust the pH of the homogenate to 7.7. The homogenate was then centrifuged at 20 000 *g* for 15 min, and phenylmethanesulfonyl fluoride was added to the supernatant to a final concentration of 0.18 mM. This material was applied to a DEAE-cellulose column (8.0 mL) equilibrated with buffer A (20 mM Tris/HCl, 2 mM EDTA, 0.5 mM dithiothreitol, pH 7.5). The column was washed with buffer A until the A_{280} of the eluate was below 0.5. The bound proteins were eluted using a 40-mL linear gradient of 0–0.3 M NaCl in buffer A at a flow rate of 0.5 $mL \cdot min^{-1}$. The fractions were assayed for protein, ionic strength and protein kinase activity.

The fractions of the first peak of protein kinase activity, eluted at 5 $mS \cdot cm^{-1}$, were pooled; ammonium sulphate was added to a concentration of 0.5 M and the solution applied to a phenyl-Sepharose column (1.6 mL) equilibrated in buffer A containing 0.5 M ammonium sulphate. It was washed with 10 column volumes of the equilibration buffer, and the bound protein kinase eluted with buffer A. Fractions of volume 0.5 mL were collected and analysed for protein kinase activity and ionic strength. The fractions with kinase activity were pooled and applied to a 100-μL Mono Q column connected to a SMART system. The column was washed until A_{280} approached zero, and the protein kinase was eluted with a linear gradient from buffer A to buffer A containing 0.3 M NaCl in a total volume of 1.5 mL. Fractions of volume 50 μL were collected and pooled according to the peak of protein kinase activity. A typical phosphorylation assay for determination of protein kinase activity consisted of 10 μL sample fraction, 10 μL 150 mM Tris/HCl/5 mM $CaCl_2$, pH 7.5, 10 μL 20 mM Tris/HCl, pH 7.5 (containing lipids when tested), 20 μL 3.75 $mg \cdot mL^{-1}$ histone or 100 μM peptide VLARLHSVRER or LLRLHSLRE and 10 μL 0.75 mM [γ-^{32}P]ATP/37.5 mM $MgCl_2$. When the stimulation of the protein kinase activity by phospholipid and diacylglycerol was tested, 60 $\mu g \cdot mL^{-1}$ phosphatidylserine and 1 $\mu g \cdot mL^{-1}$ diolein were included in the mixture. After 10 min of incubation, an aliquot of 25 μL was transferred to a 2 × 2 cm piece of Whatman 3MM paper in the case of histone or phosphocellulose paper in the case of peptides. The papers were immediately immersed in ice-cold 75 mM phosphoric acid and washed with 75 mM H_3PO_4 five times for 10 min, dried at 80 °C, and the radioactivity was measured as Çerenkov radiation.

Autophosphorylation

Fractions from the Mono Q column with CDPK activity were incubated with 0.1 mM ATP labelled with [γ-^{32}P]ATP (specific radioactivity ≈ 100 $c.p.m. \cdot pmol^{-1}$) and 5 mM $MgCl_2$ at 25 °C for 30 min. The reaction was terminated by dilution of the incubation mixture 1.5 times with SDS/PAGE sample buffer containing 65 mM Tris/HCl, pH 6.8, 10% glycerol, 2.1% SDS and 5% 2-mercaptoethanol. The proteins were separated by SDS/PAGE (10% gel) and subjected to autoradiography.

Western blotting

The protein kinase pool from Mono Q was separated by SDS/PAGE (12.5% gel) and transferred to a nitrocellulose filter using a Bio-Rad Trans-Blot®SD Semi-Dry Transfer Cell in transfer buffer containing 48 mM Tris/HCl, pH 9.2, 39 mM glycine and 20% methanol. After the transfer of the proteins to the membrane, nonspecific binding was blocked overnight at 4 °C in Tris/NaCl buffer (10 mM Tris/HCl, pH 7.4, 0.9% NaCl) containing 5% powdered milk, free of fat. Then the blot was incubated with a monoclonal antibody against the calmodulin-like domain of soybean CDPK, for 1 h at room temperature. Next, the blot was washed four times for 10 min in Tris/NaCl/Tween buffer (Tris/NaCl plus 0.1% Tween 20). The reaction with secondary antibody was preformed for 1 h at room temperature using alkaline phosphatase-conjugated rabbit antibody. After the blot had been washed four times for 10 min in Tris/NaCl/Tween buffer, the reaction was visualized using NBT/BCIP (Promega) in 10 mL 100 mM Tris/HCl, pH 9.0, containing 100 mM NaCl and 5 mM $MgCl_2$.

Purification of PKA and PKC

PKC was isolated from pig spleen by the method of Parker *et al.* [9] as modified by Ferrari *et al.* [10], except for the final chromatography on phenyl-Sepharose. PKA was purified from pig heart essentially as described by Zoller *et al.* [11].

Phosphorylation of synthetic peptides

For the maize CDPK-1, the reactions were carried out in a final volume of 150 μL containing substrate peptide at various concentrations, 44 mM Tris/HCl, pH 7.5, 0.005% Triton X-100, 0.7 mM $CaCl_2$, 0.4 mM EDTA, 0.4 $mg \cdot mL^{-1}$ BSA, 5 mM $MgCl_2$, 0.1 mM [γ-^{32}P]ATP and the protein kinase. In the inhibition experiments with the peptide inhibitor RLKQFSA-nLeu-NKLKK-nLeu-ALRVIAERL-amide (where nLeu is norleucine), derived from the inhibitory region of soybean CDPKα [12], the substrate peptide VLARLHSVRER was used at a 10 μM concentration together with 0.1 mM [γ-^{32}P]ATP. In the experiments in which stimulation by $CaCl_2$ and inhibition by EGTA were studied, 45 μM LARLHSVRER was used.

For PKC, the phosphorylation reaction was performed in a total volume of 150 μL, containing peptide substrate at various concentrations, 60 mM Tris/HCl (pH 7.5), 0.0013% Triton X-100, 0.75 mM $CaCl_2$, 60 $\mu g \cdot mL^{-1}$ phosphatidylserine, 1 $\mu g \cdot mL^{-1}$ diolein, 0.5 mM EDTA, 0.5 $mg \cdot mL^{-1}$ BSA, 5 mM $MgCl_2$, 0.1 mM [γ-^{32}P]ATP and the protein kinase.

For PKA, the phosphorylation assay was performed in a total volume of 150 μL, containing peptide substrate, 42 mM Mes, 10 mM Tris/HCl (pH 6.9), 0.001% Triton X-100, 0.04 mM EGTA, 0.5 $mg \cdot mL^{-1}$ BSA, 5 mM $MgCl_2$, 0.1 mM [γ-^{32}P]ATP and the protein kinase.

All phosphorylation reactions were carried out at 30 °C. For each peptide, the reaction time and enzyme amount were

adjusted so that the initial rates of the phosphorylation reaction could be measured. The reactions were initiated by addition of the [γ-^{32}P]ATP (specific radioactivity 50–150 c.p.m.·pmol^{-1}) and monitored by the phosphocellulose paper method as described above. It was confirmed by the method used in [13] that the difference in radioactivity for different peptides was not a consequence of deviations in their binding to the phosphocellulose paper.

Kinetic analysis of phosphorylation data

The initial rates of the phosphorylation reaction were calculated from the slopes of c.p.m. vs. time plots. Michaelis–Menten parameters were calculated using the equation:

$$v = \frac{V[S]}{K_m + [S]} \quad (1)$$

In the case of peptides of low reactivity, the substrate concentrations used were too low compared with K_m, yielding linear initial velocity vs. peptide concentration plots. For these peptides, Eqn (1) can be simplified as follows:

$$v_0 = \frac{V}{K_m}[S] = k_{II}[E][S] = k'_{II}[S] \quad (2)$$

where k'_{II} is the second-order rate constant multiplied by enzyme concentration and the value of this constant corresponds to the ratio of V/K_m. The values of k'_{II} for substrates of relatively poor reactivity were calculated from the slope of the initial velocity vs. substrate concentration plots. All calculations were made using the GRAPHPAD PRISM 2.0 software package.

Synthesis of SPOTs peptides

The SPOTs-membrane-attached peptides were synthesised utilising Fmoc chemistry essentially as described in [14] according to the instructions of the manufacturer of the SPOTs kit. After completion of the synthesis, the membrane was coated with polyvinylpyrrolidone (2 mg·mL^{-1} in NaCl/P$_i$) overnight at room temperature, followed by three washes with the same buffer (10 min each).

Phosphorylation of SPOTs peptides

The phosphorylation reactions were carried out at 30 °C essentially as described in [15]. The specific radioactivity of [γ-^{32}P]ATP was 50–150 c.p.m.·pmol^{-1}. The phosphorylation reactions were initiated by immersing the SPOTs membrane in the incubation mixture, using 100–150 μL of solution for each peptide spot. Horizontally placed test tubes were rotated around the longitudinal axis at 6 rev·min^{-1} to ensure adequate mixing. The incubation mixture contained 50 mM Tris/HCl, pH 7.5, 0.8 mM calcium acetate, 0.005% Triton X-100, 0.4 mM EDTA, 0.3 mg·mL^{-1} BSA, 5 mM $MgCl_2$, 0.1 mM dithiothreitol, 0.12 mM [γ-^{32}P]ATP and CDPK-1.

After incubation, the SPOTs membrane pieces were washed five times with 75 mM H_3PO_4 for 10 min at 4 °C, dried at room temperature and subjected to autoradiography. Equal areas in the central part of each SPOT were used for quantification of the membrane-bound radioactivity. The background level of the bound radioactivity (nonspecific binding of [γ-^{32}P]ATP) was determined by using the SPOTs-attached control peptides with Ala instead of the phosphoacceptor Ser. The amount of radioactivity, related to phosphopeptide formation, was calculated as the difference between the total binding and the background radioactivity.

RESULTS

CDPK-1 from maize seedlings

The first purification step on the DEAE-cellulose column yielded two peaks, with protein kinase activity eluted at ionic strength 5 and 15 mS·cm^{-1} (Fig. 1). Elution of the enzyme was shown by phosphorylation of histone as well as of the peptide substrate LLRLHSLRE. The first peak of the peptide and histone kinase activity was pooled, and the protein kinase, arbitrarily named CDPK-1, was subsequently purified on a phenyl-Sepharose column and by Mono Q SMART chromatography. In the two latter chromatographic steps, the enzyme behaved as a single activity peak when assayed with either peptide or histone. A summary of the purification steps is presented in Table 1. The biochemical properties and substrate specificity of the protein kinase from the second peak from DEAE-cellulose chromatography (CDPK-2) were different from those of CDPK-1 and will be further characterised.

Autophosphorylation experiments were performed for fractions in the CDPK-1 activity peak after the Mono Q chromatography step, and a single radioactive band at 59–60 kDa was visualised by SDS/PAGE (Fig. 2). Fractions 13 and 14 were pooled and used for further analysis and kinetic experiments. Western blotting of the CDPK-1 pool using antibodies against the calmodulin-like domain of soybean CDPK yielded a single band at about 59 kDa (Fig. 3). The isolated protein kinase showed full activity without added Ca^{2+}

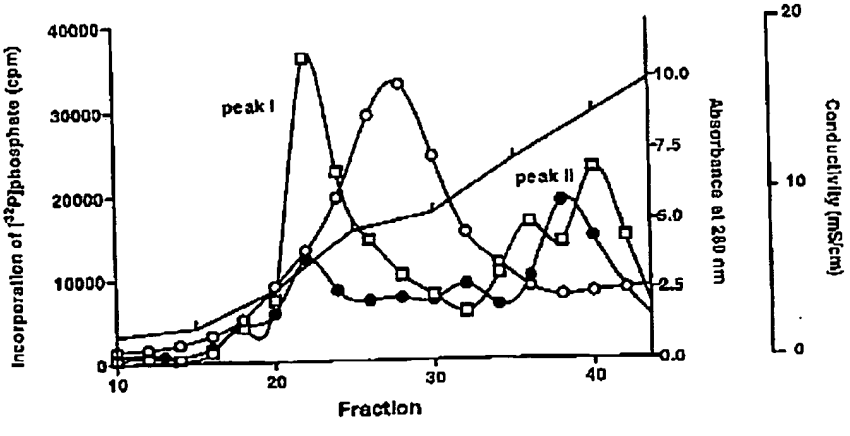


Fig. 1. Elution of protein kinase activities from the DEAE-cellulose (DE-52) column. The elution was performed using a 40-mL linear gradient of 0–0.3 M NaCl in buffer A at a flow rate of 0.5 mL·min^{-1}. (○) A_{280}; (●) phosphorylation of histone; (□) phosphorylation of peptide substrate LLRLHSVRE; (|) conductivity.

Table 1. Purification of CDPK-1 from maize seedlings. Summary of the purification steps. One absorbance unit at 280 nm was denoted as a protein concentration of 1 mg·mL^{-1}. Protein kinase activity was measured with histone as substrate.

Purification step	Protein (mg·mL^{-1})	Protein kinase activity (pmol·min^{-1}·mL^{-1})	Specific protein kinase activity (pmol·min^{-1}·mg^{-1})	Total protein kinase activity (pmol·min^{-1})	Yield (%)	Purification fold
Supernatant	94.2	100.1	1.06	2633		1
DE-52	1.4	905.5	646.8	9417	100	610
Phenyl-Sepharose	0.183	520	2842	4160	44	2681
Mono Q	0.075	4870	64 933	974	10	61 258

but was completely inhibited at 10 μM EGTA (Fig. 4A). Such a phenomenon has previously been reported for maize CDPKs and was explained by the presence of endogenous Ca^{2+} ions [16]. In the presence of 10 μM EGTA, maximal stimulation of the enzyme activity by $CaCl_2$ was achieved at concentrations of 10–30 μM (Fig. 4B), showing that the Ca^{2+} stimulation had the half effect in the low micromolar range. The results from the autophosphorylation, Western-blotting and Ca^{2+}-stimulation experiments agree with the conclusion that a single protein kinase was present in the CDPK-1 pool made up of fractions 13 and 14 from Mono Q SMART chromatography.

Stimulation of the protein kinase activity by phospholipids and diacylglycerols was tested using a phosphatidylserine/diolein mixture commonly used for activation of mammalian PKC. With CDPK-1, there was no stimulation of histone or peptide phosphorylation, whereas PKC was activated up to 25 times by lipids under similar experimental conditions.

Inhibition of CDPK-1 by the peptide inhibitor RLKQFSA-nLeu-NKLKK-nLeu-ALRVIAERL-amide derived from the inhibitory region of soybean CDPKα was studied. The IC_{50} value obtained (10.2 μM) was close to the reported K_i value (4.8 μM) of the same inhibitor for inhibition of soybean CDPKα, determined with the peptide substrate syntide-2 [12].

Phosphorylation of peptides by CDPK-1

Using the peptide sequence RVLSRLHS^{15}VRER of the phosphorylatable site of maize sucrose synthase 2, nine peptides were synthesised in order to define the minimal length required for CDPK-1 activity (Table 2). All these peptides were phosphorylated by the enzyme, and the parameters of the Michaelis–Menten rate equation were calculated for all except compound IX, which had too high a K_m value to be reached experimentally. The second-order rate constant k'_{II} was therefore calculated for this peptide from the slope of the initial velocity vs. peptide concentration plot as described in Materials and methods. The second-order rate constants for all other substrates were calculated as the ratio of V/K_m.

Two peptides (II and III in Table 2), in which both Ser residues of the parent peptide RVLSRLHSVRER were subsequently replaced by alanine, were synthesised to determine the phosphorylation site for CDPK-1. It was found that the serine residue, corresponding to Ser15 of the sucrose synthase sequence, was the only phosphate acceptor in these peptides. Because of this, all other peptides studied (IV–IX in Table 2) were synthesised with a single serine residue to avoid uncertainties in the interpretation of the data.

The second-order rate constants k'_{II} were used to compare the reactivity of different peptides. It can be seen in Table 2 that the removal of the N-terminal arginine and valine from the parent peptide did not have any dramatic effect, the constant k'_{II} decreasing less than twofold. Removal of arginine at position +4 and glutamic acid at +3 also led to a moderate decrease in k'_{II}. The greatest change in CDPK activity occurred if the leucine and alanine residues, located at positions −5 and −4, were removed. The absence of these amino acids caused a 78-fold and 22-fold decrease in k'_{II}, respectively.

Phosphorylation of CDPK-1 substrates by PKA and PKC

The peptides tested above as substrates for maize CDPK-1 were also phosphorylated by mammalian PKC and PKA (Table 2). It

Table 2. Kinetic constants of peptide phosphorylation by maize CDPK-1 and mammalian PKC and PKA. ND, not determined; NA, no activity.

	Peptide	CDPK-1			PKC			PKA		
		K_m (μM)	V (μmol·min^{-1}·mg^{-1})	$10^3 \times k'_{II}$ (L·min^{-1}·mg^{-1})	K_m (μM)	V (μmol·min^{-1}·mg^{-1})	$10^3 \times k'_{II}$ (L·min^{-1}·mg^{-1})	K_m (μM)	V (μmol·min^{-1}·mg^{-1})	$10^3 \times k'_{II}$ (L·min^{-1}·mg^{-1})
I	RVLSRLHSVRER	0.6 ± 0.2	0.97 ± 0.02	1515.6	1.90 ± 0.02	3.75 ± 0.01	1973.7	ND	ND	0.78 ± 0.07
II	RVLARLHSVRER	0.7 ± 0.2	0.94 ± 0.02	1305.6	2.8 ± 0.7	4.3 ± 0.3	1517.9	ND	ND	1.10 ± 0.02
III	RVLSRLHAVRER	NA	NA	NA	NA	NA	NA	NA	NA	NA
IV	VLARLHSVRER	1.5 ± 0.4	1.61 ± 0.07	1073.3	6.8 ± 0.6	4.4 ± 0.3	645.7	ND	ND	0.60 ± 0.02
V	LARLHSVRER	2.1 ± 0.3	1.76 ± 0.06	830.2	8.8 ± 1.0	4.3 ± 0.2	484.7	ND	ND	0.43 ± 0.02
VI	LARLHSVRE	8.7 ± 0.9	1.90 ± 0.06	217.9	41 ± 11	5.3 ± 0.6	128.2	ND	ND	0.25 ± 0.01
VII	LARLHSVR	6.6 ± 1.5	0.86 ± 0.07	129.8	141 ± 30	3.7 ± 0.4	26.6	ND	ND	0.31 ± 0.01
VIII	ARLHSVRER	176 ± 69	1.9 ± 0.4	10.7	176 ± 55	3.0 ± 0.5	17.3	173 ± 58	0.052 ± 0.009	0.3
IX	RLHSVRER	ND	ND	0.5 ± 0.1	ND	ND	7.8 ± 0.2	ND	ND	0.16 ± 0.01
	RRASVA	ND	ND	ND	ND	ND	ND	15 ± 2	7.4 ± 0.9	492
	KRAKRKTAKKR	ND	ND	ND	1.4 ± 0.7	1.7 ± 0.2	1195.8	ND	ND	ND

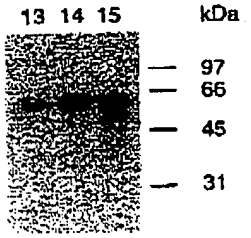


Fig. 2. Autophosphorylation of fractions 13, 14 and 15 from the CDPK activity peak eluted from the Mono Q column. The fractions were incubated with 0.1 mM [γ-^{32}P]ATP (specific radioactivity ≈ 100 c.p.m.·pmol^{-1}) and 5 mM $MgCl_2$ at 25 °C for 30 min. The reaction was terminated by the addition of SDS/PAGE sample buffer and the proteins were electrophoretically separated by SDS/PAGE (10% gel) and subjected to autoradiography.

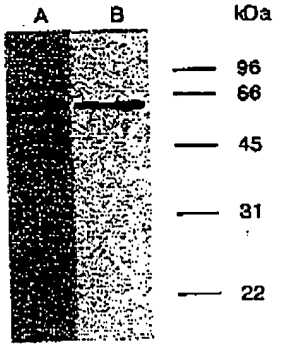


Fig. 3. Analysis of the pool of CDPK-1 from the Mono Q column by SDS/PAGE and Western blotting with antibody against the soybean CDPK. Lane A, Coomassie-stained gel after SDS/PAGE. The polyacrylamide content of the gel was 12.5%. Lane B, Western blotting. The positions of the protein markers are indicated on the right.

can be seen that the reactivity with peptides I–VI was similar for CDPK-1 and PKC. Comparison with the kinetic data for a highly effective PKC substrate KRAKRKTAKKR [15] showed that the peptides studied are also highly specific substrates for PKC (Table 2). In contrast, all the peptides tested were relatively poor substrates for PKA, showing k'_{cat} values about three orders of magnitude lower than for the conventional PKA peptide substrate RRASVA (Table 2).

Phosphorylation of SPOTs-membrane-attached peptides by CDPK-1

Using the sequence LARLHSVRER, a series of SPOTs-membrane-attached peptides were synthesised to obtain substrates with hydrophobic (Ala, Leu), cationic (Lys) and anionic (Glu) amino acids substituted at each position, while the rest of the peptide remained unchanged. Phosphorylation of these peptides by maize CDPK-1 was determined as the amount of radioactivity transferred from [γ-^{32}P]ATP to the serine residues of the attached substrates (Table 3). It was found that no more than 1% of the total amount of the attached peptides was phosphorylated under the experimental conditions used. Therefore the results obtained correspond to the initial velocity of the phosphorylation reaction. The results were normalised relative to the reactivity of the SPOTs-attached parent peptide LARLHSVRER and were used to characterise the substrate specificity of CDPK-1.

Table 3. Relative reactivity with maize CDPK of structural analogues of LARLHSVRER synthesised on SPOTs membranes. Each amino acid in this peptide was substituted subsequently with A, L, R and E at each position except for serine. In each structure, only one substitution was made thereby generating a set of 30 separate peptides. The activity of the initial peptide LARLHSVRER was taken to be equal to 1.00.

Amino acid	Amino acid position								
	−5	−4	−3	−2	−1	+1	+2	+3	+4
A	1.02	1.00	0.34	1.20	1.13	0.89	0.31	0.74	0.61
L	1.00	0.78	0.095	1.00	0.75	0.73	0.31	0.85	0.58
R	0.46	0.80	1.00	0.70	0.72	0.53	1.00	0.01	1.00
E	0.39	0.97	0.03	0.46	0.99	0.82	0.06	1.00	0.27
H	–	–	–	–	1.00	–	–	–	–
V	–	–	–	–	–	1.00	–	–	–

The results listed in Table 3 show that the positions −3, +2 and +4 were the most sensitive to replacement of Arg by other amino acids, especially by the negatively charged Glu. On the other hand, Arg was not tolerated in position +3, where Glu yielded the best phosphorylation rate. Hydrophobic residues were important in position −5, while in other positions the structural requirements were less pronounced. Thus the preliminary substrate consensus motif for maize CDPK-1 may

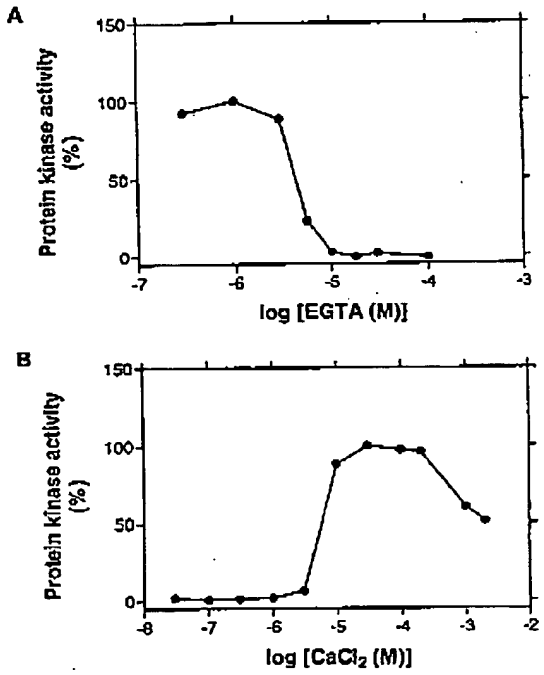


Fig. 4. Ca^{2+}-dependence of the CDPK-1 from maize seedlings. The initial velocities of the peptide LARLHSVRER phosphorylation were measured by the standard protocol described in Materials and methods, except that EDTA was not included. (A) Inhibition of the kinase activity with EGTA. Protein kinase activity in the absence of EGTA was taken as 100%. (B) Stimulation of protein kinase activity by Ca^{2+} in the presence of 10 μM EGTA. Protein kinase activity at 30 μM $CaCl_2$ was taken as 100%.

be represented by the sequence, $A/L_{-5}X_{-4}R_{-3}X_{-2}X_{-1}SX_{+1}R_{+2}Z_{+3}R_{+4}$, in which X denotes a position with no strict amino acid requirements and Z a position strictly not tolerating arginine compared with the other three varied amino acids.

DISCUSSION

The maize protein kinase isolated and studied in the present work was activated by Ca^{2+} ions in the low micromolar range, but its activity was not stimulated by the phospholipid/diacylglycerol mixture. The enzyme preparation showed both a single band of autophosphorylation and a single immunoreactive band with the monoclonal antibody raised against the calmodulin-like domain of the soybean CDPK. These data confirm that the CDPK-1 preparation contained only one protein kinase. The molecular size estimated from both autophosphorylation and blotting experiments was close to the predicted molecular mass of 59.4 kDa for a cloned CDPK from maize seedlings [17]. A molecular mass of 60 kDa has also been predicted for a cloned pollen-specific maize CDPK [18]. In addition, it was found that the activity of the isolated protein kinase was inhibited by the peptide inhibitor of soybean CDPKα. The enzyme also showed an ability to bind to a hydrophobic matrix such as phenyl-Sepharose, and its most important substrate-recognition motif involved basic residues in the phosphorylatable sequence. All these characteristics indicate that the enzyme belongs to the plant CDPK family [1].

The substrate specificity of maize CDPK-1 was indeed oriented to phosphorylation of the Ser15 residue of sucrose synthase 2, as the consensus motif, $A/L_{-5}X_{-4}R_{-3}X_{-2}X_{-1}SX_{+1}R_{+2}Z_{+3}R_{+4}$, derived from the SPOTs analysis agrees well with the peptide sequence around the phosphorylatable site of this substrate protein, RVLSRLHS[15]VRER. The other serine in this peptide (Ser11 in sucrose synthase 2) was not phosphorylated by CDPK-1, as it did not possess all the important recognition elements, except the motif RXXS.

The consensus motif for maize CDPK-1 is in agreement with results from specificity studies on CDPK isoforms from wheat germ [19], soybean [20] and a lower plant, Mougeotia [21], in which the location of basic amino acids in position −3 from the substrate phosphorylation site appears to be critical for substrate recognition. The leucine in position −5 has been shown to be an important substrate-recognition element for two forms of CDPK from spinach leaf [22]. However, the CDPK forms from these plants did not exhibit any strict requirements for basic or acidic amino acids at the C-terminal side of the serine. Thus the fragment RZR (in which Z is a hydrophobic or acidic amino acid but not a basic one) in positions from +2 to +4 seems to be specifically characteristic for maize CDPK-1. Recognition of this motif is probably based on electrostatic interactions, as the enzyme did not tolerate amino acids of the opposite charge in this region (Table 3). However, this fragment alone was not sufficient for efficient phosphorylation of peptides. For example, peptides VIII and IX, which lack the Leu in position −5 but possess the fragment RER, were both poor substrates for CDPK-1. From these observations it can be concluded that CDPK-1 has a complex and rather unique substrate specificity.

Part of the consensus motif of maize CDPK-1 resembles the consensus sequence K/RXXS/TXK/R reported for PKC [23]. Therefore it was not surprising that the second-order rate constants k'_{II} of the phosphorylation reaction of several analogues of the peptide RVLARLHSVRER exhibited a similar specificity pattern for both maize CDPK-1 and mammalian PKCβ. At the same time, PKA phosphorylated these peptides slowly and revealed a very 'flat' structure–activity relationship. The slowness of these phosphorylation reactions can be related to the presence of an arginine residue in positions +2 and +4, as PKA cannot tolerate substrates with basic residues on the C-terminal side of serine [24]. Thus the substrate specificity of maize CDPK-1 is distinct from that of PKA.

At the same time, it is interesting to note that PKA and PKC are close phylogenetic neighbours, both belonging to the AGC group of the eukaryotic protein kinase superfamily. This group is relatively distinct from CDPKs, which belong to the CaMK group of the superfamily (for nomenclature see [25]). Therefore it may be that representatives of distinct protein kinase subgroups that have similar cellular functions related to mediation of Ca^{2+} signalling have developed similar substrate-specificity patterns via evolutionarily separate pathways. The second possibility is that the peptides derived from the maize sucrose synthase phosphorylation site sequence contain the original substrate-specificity motif of a common phylogenetic ancestor of the AGC and CaMK groups. According to this explanation, CDPK and PKC have retained the specificity for this sequence during evolution, whereas the substrate specificity of PKA has diverged and developed via a different pathway. The overlapping substrate-specificity patterns of PKC and CDPK, together with the similar sensitivity of these enzymes to various inhibitors found previously [1], seem to agree with the hypothesis that, by utilising CDPKs, plants may have bypassed the need to develop kinases homologous to PKC [26]. This hypothesis is supported by the lack of any major success in purifying or cloning PKC-like enzymes from plants.

ACKNOWLEDGEMENTS

This work was supported by EU INCO COPERNICUS grant IC 15-CT96-0900, Swedish Medical Research Council grant no. K98-03X, Swedish Agricultural and Forestry Research Board grant 729.1181/97 and Estonian Science Foundation grants ETF 3348 and ETF 3916. We thank Professor Alice C. Harmon for providing the antibodies against soybean CDPK and Dr Jadwiga Szczegielniak for performing the Western-blotting experiments.

REFERENCES

1. Roberts, D.M. & Harmon, A.C. (1992) Calcium-modulated proteins: targets of intracellular calcium signals in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 43, 375–414.
2. Huber, S.C., Bachmann, M. & Huber, J.L. (1996) Post-translational regulation of nitrate reductase activity: a role for Ca^{2+} and 14–3–3 proteins. *Trends Plant Sci.* 1, 432–438.
3. Douglas, P., Pigaglio, E., Ferrer, A., Halford, N.G. & MacKintosh, C. (1997) Three spinach leaf nitrate reductase–3-hydroxy-3-metoxy-glutaryl-CoA reductase kinases that are regulated by reversible phosphorylation and/or Ca^{2+} ions. *Biochem. J.* 325, 101–109.
4. Sheen, J. (1996) Ca^{2+}-dependent protein kinases and stress signal transduction in plants. *Science* 274, 1900–1902.
5. Harper, J.F., Sussman, M.R., Schaller, G.E., Putnam-Evans, C., Charbonneau, H. & Harmon, A.C. (1991) A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. *Science* 252, 951–954.
6. Huber, S.C., Huber, J.L., Liao, P.-C., Gage, D.A., McMichael, R.W. Jr, Chourey, P.S., Hannah, L.C. & Koch, K. (1996) Phosphorylation of Serine-15 of maize leaf sucrose synthase. *Plant Physiol.* 112, 793–802.
7. Lindblom, S., Ek, P., Muszynska, G., Ek, B., Szczegielniak, J. & Engström, L. (1997) Phosphorylation of sucrose synthase from maize seedlings. *Acta Biochim. Pol.* 44, 809–818.
8. Pinna, L.A. & Ruzzene, M. (1996) How do protein kinases recognize their substrates? *Biochim. Biophys. Acta* 1314, 191–225.
9. Parker, P.J., Stabel, S. & Waterfield, M.D. (1984) Purification to

homogeneity of protein kinase C from bovine brain: identity with the phorbol ester receptor. *EMBO J.* **3**, 953–959.

10. Ferrari, S., Marchiori, F., Borin, G. & Pinna, L.A. (1985) Distinct structural requirements of Ca^{2+}/phospholipid-dependent protein kinase (protein kinase C) and cAMP-dependent protein kinase as evidenced by synthetic peptide substrates. *FEBS Lett.* **184**, 72–77.
11. Zoller, M.J., Kerlavage, A.R. & Taylor, S.S. (1979) Structural comparisons of cAMP-dependent protein kinases I and II from porcine skeletal muscle. *J. Biol. Chem.* **254**, 2408–2412.
12. Harmon, A.C., Yoo, B.-C. & McCaffery, C. (1994) Pseudosubstrate inhibition of CDPK, a protein kinase with a calmodulin-like domain. *Biochemistry* **33**, 7278–7287.
13. Toomik, R., Ekman, P. & Engström, L. (1992) A potential pitfall in protein kinase assay: phosphocellulose paper as an unreliable adsorbent of produced phosphopeptides. *Anal. Biochem.* **204**, 311–314.
14. Frank, R. (1992) Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* **48**, 9217–9232.
15. Toomik, R. & Ek, P. (1997) A potent and highly selective peptide substrate for protein kinase C assay. *Biochem. J.* **322**, 455–460.
16. Ogawa, N., Yabuta, N., Ueno, Y. & Izui, K. (1998) Characterization of a maize Ca^{2+}-dependent protein kinase phosphorylating phosphoenolpyruvate carboxylase. *Plant Cell Physiol.* **39**, 1010–1019.
17. Saijo, Y., Hata, S., Sheen, J. & Izui, K. (1997) cDNA cloning and prokaryotic expression of maize calcium-dependent protein kinases. *Biochim. Biophys. Acta* **1350**, 109–114.
18. Estruch, J.J., Kadwell, S., Merlin, E. & Crossland, L. (1994) Cloning and characterization of a maize pollen-specific calcium-dependent calmodulin-independent protein kinase. *Proc. Natl Acad. Sci. USA* **91**, 8837–8841.
19. Polya, G.M., Morrice, N. & Wettenhall, R.E.H. (1989) Substrate specificity of wheat embryo calcium-dependent protein kinase. *FEBS Lett.* **253**, 137–140.
20. Lee, J.-Y., Yoo, B.-C. & Harmon, A.C. (1998) Kinetic and calcium-binding properties of three calcium-dependent protein kinase isoenzymes from soybean. *Biochemistry* **37**, 6801–6809.
21. Roberts, D.M. (1989) Detection of a calcium-activated protein kinase in Mougeotia by using synthetic peptide substrates. *Plant Physiol.* **91**, 1613–1619.
22. Bachmann, M., Shiraishi, N., Campbell, W.H., Yoo, B.-C., Harmon, A.C. & Huber, S.C. (1996) Identification of Ser-543 as the major regulatory phosphorylation site in spinach leaf nitrate reductase. *Plant Cell* **8**, 505–517.
23. Pearson, R.B. & Kemp, B.E. (1991) Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Methods Enzymol.* **200**, 62–81.
24. Kemp, B.E., Graves, D.J., Benjamini, E. & Krebs, E.G. (1977) Role of multiple basic residues in determining the substrate specificity of cyclic AMP-dependent protein kinase. *J. Biol. Chem.* **252**, 4888–4894.
25. Hanks, S.K., Hunter, T. (1995) The eukaryotic protein kinase superfamily. In *Protein Kinase Factsbook* (Hardie, G. & Hanks, S.K., eds), pp. 7–47. Academic Press, London.
26. Stone, J.M. & Walker, J.C. (1995) Plant protein kinase families and signal transduction. *Plant Physiol.* **108**, 451–457.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 24, Issue of June 14, pp. 21630–21638, 2002
Printed in U.S.A.

Identification of Calmodulin Isoform-specific Binding Peptides from a Phage-displayed Random 22-mer Peptide Library*

Received for publication, November 12, 2001, and in revised form, March 15, 2002
Published, JBC Papers in Press, March 18, 2002, DOI 10.1074/jbc.M110803200

Ji Young Choi‡§, Sang Hyoung Lee‡§, Chan Young Park‡§¶, Won Do Heo‡¶, Jong Cheol Kim‡¶, Min Chul Kim‡¶, Woo Sik Chung‡¶, Byeong Cheol Moon¶, Yong Hwa Cheong‡¶, Cha Young Kim‡¶, Jae Hyuk Yoo‡, Ja Choon Koo‡¶, Hyun Mi Ok¶, Seung-Wook Chi‖, Seong-Eon Ryu‖, Sang Yeol Lee‡¶, Chae Oh Lim‡¶, and Moo Je Cho‡¶

From the ¶Division of Applied Life Science (BK21 Program), ‡Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Chinju 660-701 and the ‖Center for Cellular Switch Protein Structure, Korea Research Institute of Bioscience and Biotechnology, 52 Euh-eun-dong, Yusong, Daejon 305-333, Korea

Plants express numerous calmodulin (CaM) isoforms that exhibit differential activation or inhibition of CaM-dependent enzymes *in vitro*; however, their specificities toward target enzyme/protein binding are uncertain. A random peptide library displaying a 22-mer peptide on a bacteriophage surface was constructed to screen peptides that specifically bind to plant CaM isoforms (soybean calmodulin (ScaM)-1 and SCaM-4 were used in this study) in a Ca^{2+}-dependent manner. The deduced amino acid sequence analyses of the respective 80 phage clones that were independently isolated via affinity panning revealed that SCaM isoforms require distinct amino acid sequences for optimal binding. SCaM-1-binding peptides conform to a 1-5-10 ((FILVW)*XXX*(FILV)*XXXX*(FILVW)) motif (where *X* denotes any amino acid), whereas SCaM-4-binding peptide sequences conform to a 1-8-14 ((FILVW)*XXXXXX*(FAILVW)*XXXXX*(FILVW)) motif. These motifs are classified based on the positions of conserved hydrophobic residues. To examine their binding properties further, two representative peptides from each of the SCaM isoform-binding sequences were synthesized and analyzed via gel mobility shift assays, Trp fluorescent spectra analyses, and phosphodiesterase competitive inhibition experiments. The results of these studies suggest that SCaM isoforms possess different binding sequences for optimal target interaction, which therefore may provide a molecular basis for CaM isoform-specific function in plants. Furthermore, the isolated peptide sequences may serve not only as useful CaM-binding sequence references but also as potential reagents for studying CaM isoform-specific function *in vivo*.

CaM[1] is a ubiquitous intracellular Ca^{2+} receptor involved in transducing a variety of extracellular signals (1–3). In contrast to mammals, many plant species belong to a CaM multigene family that encodes various CaM isoforms (4–5). Over 30 genes encoding CaM isoforms are found in the complete nucleotide sequence of *Arabidopsis thaliana*.[2] CaMs function by modulating or regulating the activities of their target proteins (CaM-binding proteins). More than 30 CaM-binding proteins have been identified, including enzymes such as kinases, phosphatases, and nitric-oxide synthase, as well as receptors, ion channels, G-proteins, and transcription factors (7–9). Plants seem to have evolved a unique repertoire of CaM targets whose homologues in animals do not appear to be modulated by CaMs. The shear number of CaM isoforms and the diversity of CaM targets imply that these proteins in plants likely modulate a broad spectrum of processes.

CaM is dumbbell-shaped with the N- and C-terminal globular domains separated by a flexible central helix. Each lobe contains two helix-loop-helix Ca^{2+}-binding motifs referred to as "EF-hands" that are interconnected by a small β-sheet between the two Ca^{2+}-binding loops (10). Ca^{2+} binding to CaM induces conformational changes that expose hydrophobic amino acid residues on the surface of both lobes. This creates two hydrophobic pockets that are important for target peptide binding. CaM binds to a large number of proteins through interactions with specific CaM-binding domains. Therefore, one of the most intriguing questions concerning the interaction of CaM with target proteins is: how does a phylogenetically conserved protein like CaM specifically interact with so many different target sites? The determination of the three-dimensional structures of the CaM-peptide complexes greatly aided in answering this question (11–15).

Many known Ca^{2+}-dependent CaM-binding proteins possess a region that is often characterized by an amphipathic helix consisting of ~20 amino acid residues. This region contains two hydrophobic/aromatic residues that are separated by 12 intervening residues that anchor the peptide to the two lobes of CaM (2). However, sequence analyses based upon this criteria do not always identify the CaM-binding region of a protein, and therefore, the CaM-binding regions of target proteins do not always fit these criteria. Furthermore, the presence of a novel class of CaM-binding proteins that assume a non-helical conformation in the CaM complex has been suggested (16). The CaM-binding domains of these proteins are characterized by the dominance of basic amino acids in contrast to the canonical motif in which hydrophobic amino acids are dominant. An attempt was made to examine the elements common to the many reported CaM-

* This work was supported by Grant 2000-2-20900-001-1 from KOSEF, the National Research Laboratory (2000), and BK21 Program. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ These authors contributed equally to this work.

** To whom correspondence should be addressed: Division of Applied Life Science, Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Chinju 660-701, Korea. Tel.: 82-55-751-5957; Fax: 82-55-759-9363; E-mail: mjcho@nongae.gsnu.ac.kr.

[1] The abbreviations used are: CaM, calmodulin; BSA, bovine serum albumin; HRP, horseradish peroxidase; gpIII, minor coat protein; PDE, 3′,5′-cyclic nucleotide phosphodiesterase; SCaM, soybean calmodulin; TBS, Tris-buffered saline; TFE, 2,2,2-trifluoroethanol; GST, glutathione *S*-transferase; Mops, 4-morpholinepropanesulfonic acid.

[2] Munich Information Center for Protein Sequences, mips.gsf.de/proj/thal/.

	* * * * * *	
SCaM-1	DKDGDGCITTKE	
SCaM-4	DKDGDGCITVEE	EF-hand I
B-CaM	DKDGDGTITTKE	
SCaM-1	DADGNGTIDFPE	
SCaM-4	DADGNGTIEFDE	EF-hand II
B-CaM	DADGNGTIDFPE	
SCaM-1	DKDQNGFISAAE	
SCaM-4	DKDQNGYISASE	EF-hand III
B-CaM	DKDGNGYISAAE	
SCaM-1	DVDGDGQINYEE	
SCaM-4	DLDGDGQVNYEE	EF-hand IV
B-CaM	DIDGDGQVNYEE	
Consensus	DKDGDGTIDFEE	

FIG. 1. **Comparison of amino acid sequences of four EF-hands from SCaM-1, SCaM-4, and bovine brain CaM (*B-CaM*).** Consensus shows the most frequently observed amino acid residue in each position of the Ca^{2+}-binding loop region of EF hand, which is derived from the comparison of known Ca^{2+}-binding proteins (23). *Underlined* amino acid residues indicate higher than 70% conservation. Ca^{2+}-binding ligands are denoted by *asterisks*, and residues conformed to the consensus are indicated by *boldface letters*.

binding regions from different proteins. From sequence comparisons of CaM-binding peptides and CaM-binding domains within the CaM-regulated proteins, three classes of CaM-binding motifs have emerged (17). They include a modified variant of the IQ motif as a consensus for Ca^{2+}-independent binding and two related motifs for Ca^{2+}-dependent binding, termed 1-8-14 and 1-5-10 (based on the conserved hydrophobic residues within these motifs). Based on these structures and sequence information, a web-based data base is available for quickly searching for a potential CaM-binding site in a given protein sequence (18).

In plant cells, multiple CaM genes code for numerous CaM isoforms in wheat (19), potato (20), *Arabidopsis* (21), and soybean (22). We have recently cloned five CaM isoforms from soybean (SCaM-1–5). Although some of these isoforms (*i.e.* SCaM-1–3) are more than 90% identical to mammalian CaM, two (SCaM-4 and SCaM-5) exhibit only a 78% homology with SCaM-1 and are therefore the most divergent isoforms reported thus far in the plant and animal kingdoms. SCaM-4 is considered a *bona fide* CaM isoform based on the following characteristics. At primary structure level, SCaM-4 has conserved four putative EF-hands and a central linker region, hallmarking structural features of CaM (22). In addition, most of amino acid exchanges occur outside EF-hands, and the number of total amino acid residues are also conserved (148 amino acid residues for SCaM-1 and mammalian CaM, and 149 amino acid residues for SCaM-4 (22)). When compared with the preferred consensus amino acid residues of EF-hands derived from known Ca^{2+}-binding proteins (23), the residues in all four Ca^{2+}-binding loops of SCaM-4 conform to the consensus (Fig. 1), suggesting that SCaM-4 can bind four Ca^{2+} molecules. All SCaM isoforms including SCaM-4 are ubiquitously expressed in various plant tissues and show similar subcellular localization patterns to the highly conserved SCaM-1 (22, 24). Intriguingly, the cellular level of SCaM-4 rapidly rises upon specific stimuli such as pathogen infection in the same cells constitutively expressing SCaM-1 (25). Furthermore, SCaM-4 has the ability to modulate activity of many CaM-dependent enzymes. However, SCaM-4 has distinguished ability from SCaM-1 in the activation of CaM target enzymes, which can be categorized into three different types (22, 26–29) as follows: 1) enzymes activated equally by both SCaM-1 and SCaM-4 (*e.g.* phosphodiesterase (PDE), plant Ca^{2+}-ATPase, plant glutamate decarboxylase, and CaM-dependent protein kinase II); 2) enzymes activated only by SCaM-1 (*e.g.* calcineurin, myosin light chain kinase, red cell Ca^{2+}-ATPase, and plant NAD kinase); and 3) enzymes activated only by SCaM-4 (*e.g.* nitric-oxide synthase). SCaM-1 and SCaM-4 also exhibit differences in their Ca^{2+} concentration requirements for target enzyme activation (28). Studies defining CaM-binding sequences and analyzing their interactions with CaM have shed light on the specificity of the interactions between CaM and the particular target molecules.

Here we have adapted an approach for defining the SCaM-1 and SCaM-4-binding peptide sequences using a random peptide bacteriophage display library. Bacteriophage display and affinity selection of phage-displayed peptide libraries, a technique based on screening a library of foreign peptides displayed on the surface of M13 bacteriophage, has proven to be a very useful tool for characterizing a number of protein-protein interactions. Due to the physical linkage of the expressed peptide with its genetic sequence, libraries numbering from 10^8 to 10^{10} peptides have been rapidly screened for a wide variety of applications. This useful tool has been used to map antibody epitopes and to discover peptide ligands for membrane receptors and cytosolic proteins in recent years (30). In this work, we have searched SCaM isoform (SCaM-1 and -4)-favored peptide sequences from a phage display library, and we defined novel SCaM-1- and SCaM-4-specific binding sequences. Furthermore, we have analyzed the interactions between SCaM-1 and SCaM-4 and their respective binding peptides using a variety of biochemical techniques including gel overlay, gel mobility shift, fluorescence spectroscopy, and phosphodiesterase competition assays. Our data show that SCaM isoforms possess different sequences for optimal binding to specific targets. This may provide a molecular basis for CaM isoform-specific function in plants.

EXPERIMENTAL PROCEDURES

Construction of a Random 22-mer Phage Library—The phagemid vector pCANTAB5E (Amersham Biosciences) was modified to construct a vector for our library (Fig. 2*A*). Two oligonucleotides (5′-CCGGCCGGCGATATCAGCGGC-3′ and 5′-GGCCGCGCCGCTGATATCGCCGGCCGGCT-3′) were mixed, heated to 75 °C for 10 min, and allowed to hybridize by cooling slowly at room temperature for 30 min. The hybrids were inserted at the *NotI* and *SfiI* site of pCANTAB5E to introduce *NotI* and *EcoRV* sites and a sequence coding for junction amino acid residues. The resulting phage peptide library vector was named pCANTAB5F. These vectors were digested with *NotI* and *EcoRV* to clone double-stranded DNA coding for random peptides. This process is shown in Fig. 2*B*. Single-stranded DNA encoding 22-residue random peptides (5′-$(NNK)_{22}$GCGGCCGCAGGTGCGC-3′) and one oligonucleotide (5′-GCGCACCTGCGGCCGC-3′), in which N is A, C, G, or T (equimolar), and K is C or T (equimolar) were hybridized and converted into double-stranded DNA using Klenow DNA polymerase. The DNA was cut by *NotI* to produce one blunt end and one *NotI* site and then ligated into pCANTAN5F. The resulting DNA sequence and amino acid sequence of the N-terminal region of gpIII are shown in Fig. 2*C*. Twenty electroporations, each using 0.5 μg of the DNA constructed as above and 100 μl of electrocompetent *Escherichia coli* TG1 cells, yielded 4.2×10^8 transformants producing infectious phages. The helper phage M13K07 was used to produce a phage library from the bacteria harboring library phagemids.

Isolation of SCaM Isoform Binding Phages by Biopanning—The library was screened either with SCaM-1 or SCaM-4 isoforms (22) coated on 35 × 10-mm plastic Petri dishes. These Petri dishes were coated with SCaM isoforms by incubating them with 1 ml of the SCaM protein (100 μg/ml in TBS) for 1 h at 37 °C. The residual binding capacity of the dish was blocked with 0.1% BSA. About 2.3×10^{11} colony-forming units of the phage library in TBS (150 mM NaCl in 50 mM Tris-HCl, pH 7.5),

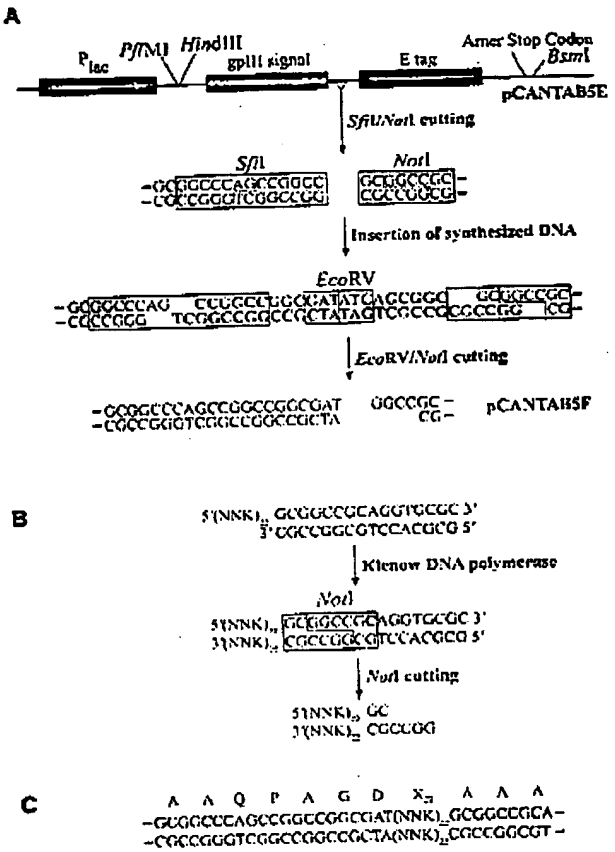


FIG. 2. **Construction of the random peptide library.** Recognition sequences and restriction sites of enzymes are indicated by *boxes* and *edged lines*, respectively. The *letter N* stands for an equal mixture of A, G, C, and T, and *K* stands for an equal mixture of Cys and Thr. *A*, procedure is described for construction of the peptide library vector pCANTAB5F. The amino acid sequence at the N terminus of gpIII and the corresponding DNA sequences are shown. *Boldface letters* correspond to the substituted or inserted residues. *B*, procedure is depicted for construction of double-stranded DNA coding for random peptides. *C*, amino acid sequence is listed for the random peptide, indicated as X_{22}. The corresponding DNA sequence is also shown.

0.1% Tween 20 were preincubated with 0.1% BSA for 1 h at room temperature to remove BSA-binding phages and were bound to the SCaM-coated dishes by incubating them overnight at 4 °C. The plates were washed ten times for 5 min with 1 ml of TBS, 0.1% Tween 20 to remove unbound phages. Phages bound to SCaM-1 or SCaM-4 were eluted in TBS, 0.1% Tween 20 with 2 mM EGTA. The eluted phages were amplified by infecting logarithmic phase *E. coli* TG1. The helper phage M13K07 was introduced into infected cells to produce phages. The resulting phages were subjected to an additional two rounds of biopanning. SCaM-1 or SCaM-4-binding phages obtained by three rounds of the selection process were plated with *E. coli* TG1 to raise colonies. Each clone was grown in 1 ml of 2× YT medium supplemented with 100 μg/ml ampicillin at 37 °C, shaking to a late log phase. After adding the Helper phage M13K07, the culture was further incubated overnight, and a clear phage supernatant was collected following a 5-min centrifugation. Each of the phages was transferred into a 96-well microtitration plate coated with the SCaM-1 and -4 isoforms as in the above biopanning method. These were tested for the binding affinities of the selected phages to SCaM-1 and -4 by enzyme-linked immunosorbent assay using an HRP-conjugated anti-M13 antibody (Amersham Biosciences). The reaction of HRP with 2',2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) was quantified by measuring the absorbance at 410 nm in a microplate spectrophotometer.

Expression of Recombinant GST Peptide Fusion Proteins in E. coli—Inserts of selected clones were amplified by PCR using modified primers both upstream and downstream of the gene III sequence (5'- TTAGCCATGGCTTTCTATGCGGCCCAG-3' and 5'-AGCGGACCCGGGTACGGCACCGGCCGA-3'), where the underlined nucleotides represent the *Sma*I and *Nco*I recognition sequence. PCR was performed using a temperature program consisting of 25 cycles of 1 min at 95 °C, 1 min at 55 °C, and 1 min at 72 °C, followed by one cycle of 5 min at 72 °C. After amplification, the PCR products were digested with *Sma*I and *Nco*I and cloned in *Sma*I- and *Nco*I-cleaved GST fusion vector, pGEX-KG (31). The ligated DNAs were used to transform *E. coli* BL21 cells. DNA was prepared as above from the cultures of individual colonies, and recombinants were identified by digestion with *Sma*I and *Nco*I and confirmed by DNA sequence analysis. Recombinant DNA techniques including transformations, plasmid preparations, gel electrophoresis, ligations, and enzyme digestions were all carried out according to Sambrook *et al.* (32). Purification of DNA from agarose gels was performed with a QIAEX II purification kit (Qiagen).

For the preparation of GST fusion proteins, transformants were grown in liquid cultures with ampicillin selection to an optical density of 0.5 at 600 nm. 1 mM isopropyl-1-thio-β-D-galactopyranoside was then added, and the cultures were further incubated for 16 h at 37 °C. Cells were collected by centrifugation and stored at −20 °C or less until used. Frozen cells were thawed on ice and resuspended in buffer M (20 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 5 mM 2-mercaptoethanol, 2 mM phenylmethylsulfonyl urea, 0.1 mM *p*-aminobenzamidine), and the cell suspensions were sonicated for 1 min at a 20% pulse. After clearing the lysate by centrifugation at 15,000 rpm for 10 min at 4 °C, the supernatant (designated as the crude CaM-PEP preparation) was used for the binding assays.

CaM Gel Overlay Assay—To prepare horseradish peroxidase (HRP)-conjugated CaMs, we conjugated SCaMs with a maleimide-activated HRP using the EZ-Link maleimide-activated HRP conjugation kit (Pierce) according to the manufacturer's instructions. Before conjugation, SCaMs were incubated in 50 mM Hepes, pH 7.0, 0.1 M dithiothreitol at 55 °C for 1.5 h to reduce cysteine residues. The reduced SCaMs were washed and concentrated in degassed phosphate-buffered saline buffer using Centricon C10s (Amicon). After conjugation, the efficiency of SCaMs conjugation to HRP was determined by SDS-PAGE. Conjugation efficiency was usually greater than 90%. Unconjugated residual CaM was removed by gel filtration. CaM gel overlays were performed as described previously (33). Ten μg of GST fusion protein was separated on an 11% SDS-polyacrylamide gel.

DNA Sequencing—DNA sequencing was carried out using primers (5'-CCATGATTACGCCAAGCTTTGGAGCC-3' and 5'-GTAAATGAATTTTCTGTATGAAG-3') and an Applied Biosystems automatic sequencer (PerkinElmer Life Sciences).

Peptide Synthesis—Peptides (Pep A, APAHALFHWGVLGSLIRLVFLS, and Pep B, CNRLLLRSLRYWGYVVLSALRL) were synthesized at the Peptide Synthesis Facility (Peptron, DaeJun, South Korea) using an Applied Biosystems model instrument. Peptides were desalted and purified by reverse-phase high pressure liquid chromatography using a C_{18} column, and their amino acid sequences were analyzed.

Gel Mobility Shift Assays—The abilities of the synthetic peptides (*i.e.* Pep A and B) to bind to the SCaM-1 and SCaM-4 isoforms were examined by the relative mobility shifts of the CaM-peptide complexes using 4 M urea polyacrylamide gel electrophoresis in the presence of the peptide and 0.1 mM $CaCl_2$ or 2 mM EGTA. Urea gels contained 13% acrylamide, 4 M urea, 0.375 M Tris-HCl, pH 8.8, and 0.1 mM $CaCl_2$, or 2 mM EGTA and were run at a constant voltage of 100 V in electrode buffer consisting of 25 mM Tris-HCl, 192 mM glycine, pH 8.3, and 0.1 mM $CaCl_2$, or 2 mM EGTA. The SCaM-1 or SCaM-4 isoform and increasing concentrations of Pep A and B (molar ratio: 0.25, 0.5, 0.75, 1.0, 1.5, 2.0, and 2.5) were incubated at room temperature for 1 h in 100 mM Tris-HCl, pH 7.2, 4 M urea, and 0.1 mM $CaCl_2$, or 2 mM EGTA. Glycerol (50%) with tracer bromphenol blue was added before the samples were loaded onto the gel.

Fluorescence Spectroscopy—The changes in the microenvironments of the peptide Trp residues upon binding to SCaM were monitored by fluorescence spectroscopy using a PerkinElmer Life Sciences LS50 luminescence spectrometer. The excitation and emission slit widths were 2 nm, and emission spectra scanning was carried out at 10 nm/min with a 1-cm path length cuvette. The Trp residues of Pep A and B incubated in Ca^{2+} buffer (50 mM Mops, pH 7.5, 0.1 M KCl, 0.1 mM $CaCl_2$) or in EGTA buffer (50 mM Mops, pH .5, 0.1 M KCl, 2 mM EGTA) at room temperature for 30 min were excited at a 295 nm excitation wavelength, and the fluorescence emission spectra in the range 295–560 nm were recorded. An excitation wavelength of 295 nm was used to decrease Tyr fluorescence in SCaM. SCaM was added to the same cuvette from a highly concentrated stock solution to maximize the dilution effects.

Circular Dichroism—CD spectra were obtained on a Jasco J-720 spectropolarimeter using a 1-mm cell at 25 °C. The peptide concentration used for CD measurements was 50 μM, and the solutions included 10 mM Tris-HCl, pH 7.5, and 50% 2,2,2-trifluoroethanol (TFE) (v/v). Secondary structure contents were calculated using the Circular Dichroism Deconvolution program (34).

Phosphodiesterase (PDE) Competition Assay—Cyclic nucleotide PDE assays were performed using commercially available bovine heart CaM-deficient PDE (Sigma). The initial 100-μl reaction volume contained buffer (100 mM imidazole HCl, 2.56 mM cAMP, 5.13 mM $MgSO_4$, 1.28 mM $CaCl_2$) and varying concentrations of SCaMs (5–200 nM) in the presence (100 nM of either Pep A or Pep B) or absence of peptides. The reaction was started by the addition of PDE (0.3 milliunits/μl). The basal level of enzyme activity was determined in the absence of SCaM, and stimulated activity was determined in the presence of either SCaM-1 or SCaM-4 and $CaCl_2$. After an incubation at 37 °C for 30 min, the reactions were stopped by placing the reaction tubes in a boiling water bath for 5 min and then on ice for 2 min. Following a brief centrifugation step, 50 μl of alkaline phosphatase (10 units) was added, and the samples were incubated for 10 min at 37 °C. The reactions were stopped by adding 500 μl of 10% trichloroacetic acid. After vortexing, the precipitates were pelleted, and the supernatants (400 μl) were transferred to new tubes. One ml of phosphate reagent (22) was then added, and samples were incubated for 15 min at 37 °C and assayed for P_i content at OD_{660}. The dissociation constants (K_d) of SCaM-1 and -4 for Pep A and B were calculated from the concentration of SCaM (nM) required to obtain half-maximal PDE activity either in the presence (100 nM) or absence of the peptides. The following equation (35) was used to calculate the dissociation constants: $K_d = ([P_t] + K - [CaM])K/([CaM] - K)$, where $[P_t]$ represents the total concentration of peptide added, and [CaM] and K are the concentrations of CaM required to obtain half-maximal activation of PDE in the presence or absence of peptides, respectively. The K_d for Pep A or B binding to SCaM-1 or -4 was calculated using the following values: $[P_t] = 100$ nM, $K = 9.7$ nM (for Pep A) or 10.6 nM (for Pep B) (calculated from the SCaM-1 or -4 activation curves in the absence of Pep A or B), [SCaM] = 33.6 nM (for Pep A) or 37.4 nM (for Pep B) (calculated from the SCaM activation curve in the presence of 100 nM Pep A or B).

RESULTS

Construction and Characterization of the Random 22-mer Peptide Phage Display Library—The 22-mer phage display random peptide libraries were constructed using a modified pCANTAB vector, as described in Fig. 2. Random peptides of 22 amino acids in length are expressed at the N terminus of the M13 protein III bacteriophage. Initial transformation of the library produced 4.8×10^8 independent clones and subsequent amplification yielded 2.3×10^{11} pfu/ml. To confirm the diversity of peptide sequences in the library, we randomly picked 80 individual phage clones and determined nucleotide sequences of random peptide regions. In Fig. 3*A*, the distribution of amino acids in the 22-mer random peptide library is reported as a percentage of the total. All amino acids were uniformly distributed in this library, taking into account the bias inherent in the codon usage ratio. The distribution of amino acid types that appears at each residue position of the sequenced inserts is shown graphically in Fig. 3*B*. Acidic (Asp and Glu), basic (Lys, Arg, and His), polar (Ser, Thr, Cys, Tyr, Asn, and Gln), and nonpolar amino acids (Gly, Ala, Val, Leu, Ile, Pro, Phe, Trp, and Met) are relatively proportional to their presence in the genetic code and generate not only linear sequence diversity but also the biochemical diversity desired in a library for screening.

Isolation of SCaM-1- and SCaM-4-binding Peptides—In an effort to study the binding sequence preferences of the SCaM-1 and SCaM-4 isoforms, we screened a phage display peptide library by biopanning with purified SCaM-1 and SCaM-4 proteins, respectively (see under "Experimental Procedures"). Briefly, SCaM-1 or SCaM-4 proteins immobilized on the surface of plastic Petri dishes were incubated with phages in the presence of Ca^{2+}. Bound phages were eluted with EGTA after extensive washing. After three rounds of selection, isolated phage clones were tested for specific binding to SCaM-1 or SCaM-4 via an enzyme-linked immunosorbent assay-based assay (see "Experimental Procedures"). Approximately 90% of the isolates exhibited strong SCaM-1- or -4-binding activity (data not shown). Each of the 80 independent phages reacting strongly in the enzyme-linked immunosorbent assay were selected from the SCaM-1- and SCaM-4 biopanning, and their nucleotide sequences were determined. In Fig. 4, their deduced amino acid sequences are shown in random peptide regions, which are aligned to reveal consensus sequences. For convenience, we designated SCaM-1-binding peptide sequences as the α-series and SCaM-4-binding peptides as the β-series. Despite the considerable sequence diversity observed among these clones, their careful alignment yields two distinct binding motifs. Based on the positions of conserved hydrophobic residues, SCaM-1-binding peptides can be fit into a 1-5-10 motif, possessing a consensus sequence of ((FILVW)*XXX*(FILV)) (where *X* denotes any amino acid). In contrast, SCaM-4-binding peptides conformed to a 1-8-14 motif of ((FILVW)*XXXXXX*(FAILVW)*XXXXX*(FILVW)). In both cases, most of these peptide sequences contain one or more basic amino acid residues, which results in an overall positive net charge for these peptides.

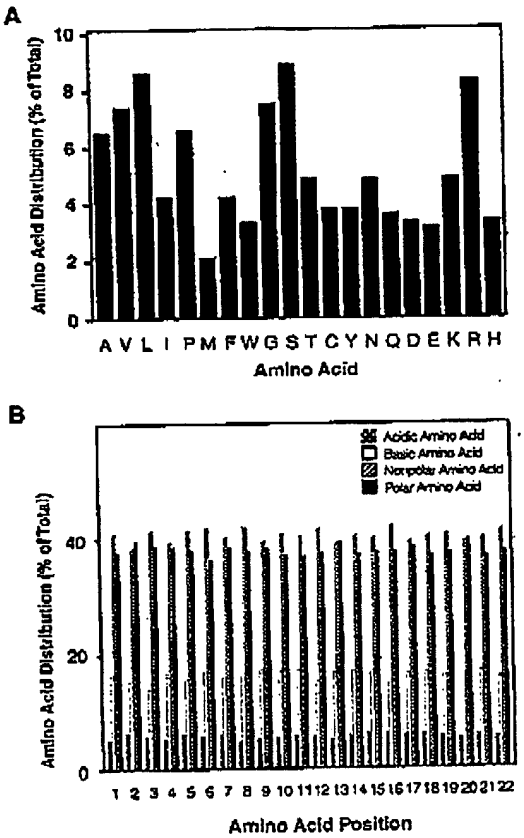


FIG. 3. **Amino acid distribution in the 22-mer random peptide library.** *A*, a total of 80 peptide sequences from the 22-mer random library was obtained and analyzed as described under "Experimental Procedures." The percentage of each total amino acid was determined. *B*, frequency and positional distribution are listed for the amino acid types in the 22-mer random peptide library. The amino acids were divided into acidic (Asp and Glu), basic (Lys, Arg, and His), nonpolar (Gly, Ala, Val, Leu, Ile, Pro, Phe, Trp, and Met), and polar (Ser, Thr, Cys, Tyr, Asn, and Gln) species. The frequency of each amino acid type at each residue position is indicated as follows: acidic, *dotted bars*; basic, *open bars*; nonpolar, *hatched bars*; and polar, *filled bars*.

A. SCaM-1 Binding Peptides:

Consensus sequence: [(FILVW)xxx(FILV)xxxx(FILVW)]

	1 5 10	net charge	n
α1	APAHALFHWGVLGSLIRLVFLS	0	13
α2	LPARFHWGVLGSLIYLRPVMLS	0	9
α3	APAHFIKWLASLGPVYLPRTSL	+1	8
α4	LGVMGQSVIKVLVKTPFWYLPG	+2	7
α5	QFPYFLMLMIKRIGKYVLAWAL	+3	7
α6	YVPMTSGVGPCQCPVGKLNWRF	+1	7
α7	LLVMRNSWALIVSQRPSGSLGW	+1	6
α8	SLLNLVVKGARFWRTFPGSFGW	+2	6
α9	HGKMRVRALWPVLRLACLGGSC	+1	4
α10	VARKEAGVWPRIRGYFSSCAFT	+1	3
α11	GPLWNGPGVKTRVVSRQLFLCA	+2	3
α12	VSLYPROSVAFVRWLFEVLCVG	+2	2
α13	RLASGSSLCKLALKPRLVGPVA	+2	1
α14	VSKLFDFHLSLVGLVASYVTWG	-1	1
α15	WGMMRVRALNPVLRLACLGGSC	+1	1
α16	VRSRGNVGSKVGLCISCSRIVF	+1	1
α17	VKASRGPQVSIPSQAVTLLRLS	0	1

B. SCaM-4 Binding Peptides:

Consensus sequence: [(FILVW)xxxxxx(FAILVW)xxxxx(FILVW)]

	1 8 14	net charge	n
β1	CNRLLLRSLRYWGYVVLSALRL	+2	11
β2	RLLLKSLRYWGHMVLSGLRGKT	+2	10
β3	IFCWGNVVSGGLKFAYVLQMVL	+1	9
β4	QASLLFWKLQIYRGLRFSLSFM	+3	9
β5	SKRLWLSLGPMLGQLLKFIFGI	+1	9
β6	GQMPPRLYLTWAASSRMAVCCG	+1	8
β7	VGSWSVPQVRSWFLEVRKLPSG	+3	7
β8	VARSYLRSLSIHARCHFRVVCL	+3	6
β9	QSRSLPRKPGSVVHAMCLRWLL	+2	4
β10	HAGHAIYSVWRLWSKRVQNRVR	+3	1
β11	QILTWSRQPVRQGPFGVAGVRL	+2	1
β12	IMSYLLRPRKLMVQVPGSVCGW	+3	1
β13	FQGGPVLLHKQLRCPVDSSLFC	+2	1
β14	SLGWFWSCTCKSLCLYSRFWRV	+2	1
β15	VRLRAPRDVPRGRASVTRAGCA	+3	1
β16	SQAIKLDRLWPYVRRIWLFLFT	+2	1

FIG. 4. **Deduced amino acid sequences derived from the phage display library by affinity selection with SCaM-1 (*A*) or SCaM-4 (*B*).** Phages from a 22-mer random peptide display library were subjected to three rounds of selection on SCaM-1 or SCaM-4 affinity panning. Each of the 80 independent phages selected for its strong enzyme-linked immunosorbent assay signal was selected from SCaM-1- or SCaM-4 biopanning. Their nucleotide sequences were determined, and the deduced amino acids are compared. *Boldface residues* indicate the three hydrophobic amino acids that are grouped into one of two types (the 1-5-10 ((FILVW)*XXX*(FILV)*XXXX*(FILVW)) motif represents the SCaM-1-binding peptides, and the 1-8-14 ((FILVW)*XXXXXX*(FAILVW)*XXXXX*(FILVW)) motif represents the SCaM-4-binding peptides) based on the positions of conserved hydrophobic residues. The frequency of appearance (*n*) and electrostatic charge of each of the isolates is indicated as shown.

Identification of SCaM-1- or -4-specific Peptides by Gel Overlay Assay—To verify the specificity of SCaM-binding phage isolation using biopanning, a CaM gel overlay assay was performed (24). First of all, three of the most frequently isolated peptides obtained from each of the SCaM-binding phages (*i.e.* α1, α2, and α3 for SCaM-1 and β1, β2, and β3 for SCaM-4) were expressed as GST fusion proteins using the vector pGEX-KG (30). As shown in Fig. 5, these GST fusion proteins exhibited molecular masses of ~29 kDa by SDS-PAGE, consistent with the predicted values of ~3 kDa for the SCaM peptide plus ~26 kDa for the GST protein. In the gel overlay assays, all α peptide fusion proteins exhibited positive signals with SCaM-1-HRP (Fig. 5*A*), whereas SCaM-4-HRP resulted in no signal. Conversely, for the β peptide fusion proteins, only SCaM-4-HRP

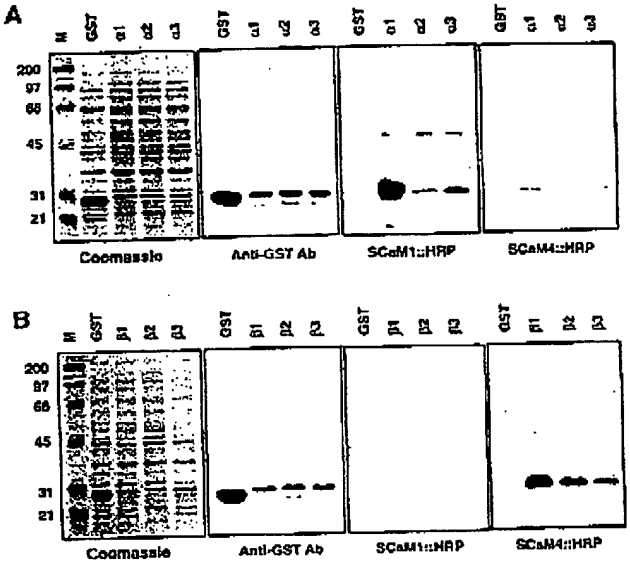


FIG. 5. **SCaM gel overlay of isolated peptides.** The GST fusion proteins were prepared as described under "Experimental Procedures." Equal amounts (10 μg) of total *E. coli* protein extracts were subjected to SDS-PAGE (11% gel), and stained by Coomassie Brilliant Blue (*left 1st panels*). The protein gels were transferred to nitrocellulose membranes and probed with anti-GST antibody (*Ab*) to detect the expression of GST fusion peptides (*2nd panels* from *left*) and probed further with either SCaM-1-HRP (*3rd panels* from *left*) or SCaM-4-HRP (*right panels*). The bound SCaM-1-HRP or SCaM-4-HRP was detected using an ECL system (Amersham Biosciences). *M* indicates protein molecular weight marker, *A* shows the SCaM-1-favored peptides, and *B* shows the SCaM-4-favored peptides.

exhibited positive binding. Interestingly, CaM binding signals for these fusion proteins correlated with their frequency of isolation in the biopanning. For example, α1 and β1, the most frequently isolated sequences for SCaM-1 and SCaM-4 binding, respectively, produced the strongest signals among the tested α and β peptides in the gel overlays (Fig. 5, *A* and *B*). In either case, the negative control (*i.e.* GST protein alone) exhibited no binding to either of the SCaM-HRP probes. In addition, no significant signal was detected when the gel overlay assays were performed in the presence of an EGTA-containing buffer (data not shown). Thus, these results indicate that these peptide sequences specifically bind to the corresponding SCaM-1 or SCaM-4 in a Ca^{2+}-dependent manner.

Gel Mobility Shift Assays—The gel overlay assays confirmed the binding of SCaM to the isolated peptide sequences. However, they did not provide quantitative information for these interactions. To examine further the SCaM-binding specificity of the isolated peptide sequences, the most frequently isolated peptide sequence was selected from each SCaM-1 or SCaM-4-binding peptide sequence. These are designated as Pep A (APAHALFHWGVLGSLIRLVFLS) and Pep B (CNRLLLRSLRYWGYVVLSALRL), respectively. The synthesized peptides were then tested by gel mobility shift to assess their interactions with the SCaM isoforms. Complexes formed between the SCaM isoforms and either Pep A or Pep B were confirmed by gel electrophoresis in the presence of 4 M urea. Urea was used since it dissociates low affinity and nonspecific complexes, leaving only the higher affinity and higher specificity complexes behind. As the ratio of peptide to Ca^{2+}-SCaM increased, we observed a band shift due to the formation of a complex between the peptide and Ca^{2+}-SCaM (Fig. 6, *A* and *D*). In the absence of peptide, SCaM migrates as a single band (data not shown). When the ratio of peptide to SCaM was equal, a mo-

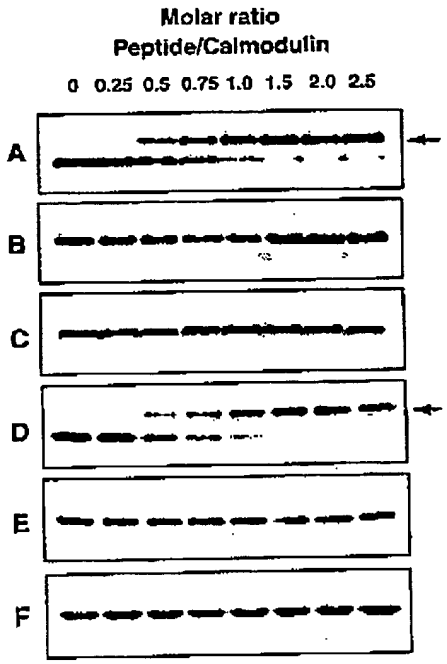


FIG. 6. Gel mobility shift assay for peptide-SCaM isoform complexes. Complex formation between SCaM isoforms and either Pep A or Pep B was determined in the presence of either 0.1 mM $CaCl_2$ or 2 mM EGTA. Purified SCaM-1 or SCaM-4 was incubated with increasing amounts of Pep A or B (Pep A or B *versus* SCaM-1 or -4 molar ratios are indicated), and then samples were separated by nondenaturing PAGE in the presence of 4 M urea. The gels were stained with Coomassie Brilliant Blue. *A* indicates SCaM-1 + Pep A + Ca^{2+}; *B* indicates SCaM-1 + Pep A + EGTA; *C* indicates SCaM-4 + Pep A + Ca^{2+}; *D* indicates SCaM-4 + Pep B + Ca^{2+}; *E* indicates SCaM-4 + Pep B + EGTA; and *F* indicates SCaM-1 + Pep B + Ca^{2+}. The *arrows* in *A* and *D* indicate the SCaM-1-Pep A or SCaM-4-Pep B complexes, respectively. Free peptide is not apparent because it had run out of the gels by the time electrophoresis was stopped.

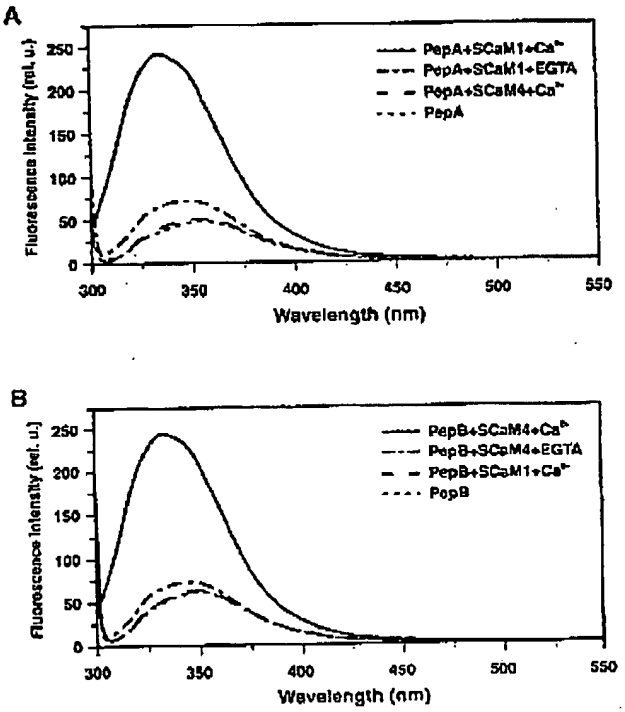


FIG. 7. Trp fluorescence spectra of peptide binding to SCaM isoforms. R_p fluorescence spectra were measured in the presence of Ca^{2+} or EGTA using a PerkinElmer Life Sciences LS50 luminescence spectrometer. An excitation wavelength of 295 nm was used to reduce Trp fluorescence from SCaM-1 (*A*) and SCaM-4 (*B*). SCaM-1 or -4 (2.5 μM) was added in a solution containing 1.9 μM Pep A or B in the presence of either 0.1 mM Ca^{2+} or 2 mM EGTA.

bility shift was observed for nearly all of the SCaM proteins. At a peptide to SCaM molar ratio of 1.5, virtually no free SCaM was detected. Taken together, these observations indicate that the peptides bind to the Ca^{2+}-bound SCaM isoforms at a 1:1 molar stoichiometry. This is consistent with the previous observation (36) that most of the well characterized CaM-binding peptides, including peptides derived from myosin light chain kinase, constitutive nitric-oxide synthase, and CaM-dependent protein kinase I form a 1:1 complex with Ca^{2+}-CaM. No mobility shift was observed when 5 mM EGTA was added to the samples, indicating that the interactions between the SCaM isoforms and the peptides are Ca^{2+}-dependent (Fig. 6, *B* and *E*). Interestingly, we could detect neither a SCaM-1-Pep B nor a SCaM-4-Pep A complex in the mobility shift assay (Fig. 6, *C* and *F*). These results suggest that in the presence of Ca^{2+}, Pep A and Pep B form a 1:1 complex with SCaM-1 and SCaM-4, respectively. Additionally, these peptides possess a certain degree of specificity toward the SCaM isoforms.

Trp Fluorescence Emission Spectra—The peptide interactions with SCaM-1 or -4 were also examined by taking advantage of the intrinsic fluorescence of the Trp residue in the peptide sequences. Fluorescence spectroscopy is a convenient method for determining the SCaM-peptide complex. The peptides have excitation and emission wavelengths of 295 and 295–560 nm, respectively. The maximum emission of the peptide in the absence of CaM occurred at 353 nm (Fig. 7). Upon addition of SCaM-1 or -4 to the reaction mixtures containing Pep A or Pep B in the presence of Ca^{2+}, the maximum fluorescence emission of the synthetic peptides shifted to shorter wavelengths (*i.e.* for SCaM-1 it was from 353 to 333 nm and for SCaM-4 it was from 353 to 335 nm) and exhibited ~5-fold increase in intensity. This relatively large blue shift indicates that the Trp residues were in a hydrophobic environment such as the interior of SCaM, which is typical for CaM-binding peptides (37). The addition of 2 mM EGTA to the same reaction mixture completely reversed the shift in the emission spectrum (Fig. 7). These results revealed that each of the SCaM peptides bind directly to its corresponding SCaM isoform.

Circular Dichroism Spectra—To examine the molecular basis of the binding specificity for the plant CaM isoforms, we characterized the secondary structures of the model peptides. Because shorter peptides tend to exist as flexible coils in pure water, we determined the secondary structures of the model peptides in the presence of 50% TFE, a compound that stabilizes secondary structures (38). In Fig. 8, the CD spectra for Pep A and Pep B obtained at 25 °C are presented. As can be seen, the spectra are very similar to each other. Both contain a minimum at 206–208 nm with a shoulder at around 222 nm, indicating the dominant presence of α-helices. From the CD spectra, the α-helix contents of Pep A and Pep B are estimated to be 72 and 76%, respectively. Thus, both Pep A and Pep B are likely to bind CaM isoforms as helices. In addition, in their SCaM-1-bound forms, both peptides showed increases in their α-helical content, further supporting this theory (data not shown).

Peptide-dependent Inhibition of CaM-activated PDE Activity—Finally, the relative affinities of Pep A and Pep B for the SCaM isoforms were determined by using a competition assay for PDE, an enzyme that is activated equally by both SCaM-1

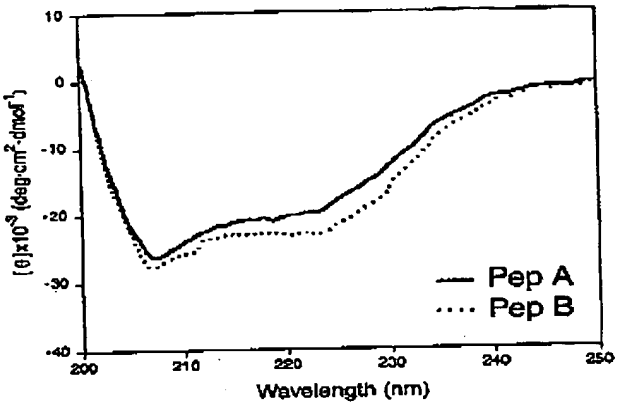


FIG. 8. **CD spectra of Pep A and Pep B.** CD spectra of Pep A (—) and Pep B (—) were measured in 50% TFE (v/v). Five consecutive scans (250–200 nm) for each sample were averaged to generate a spectrum.

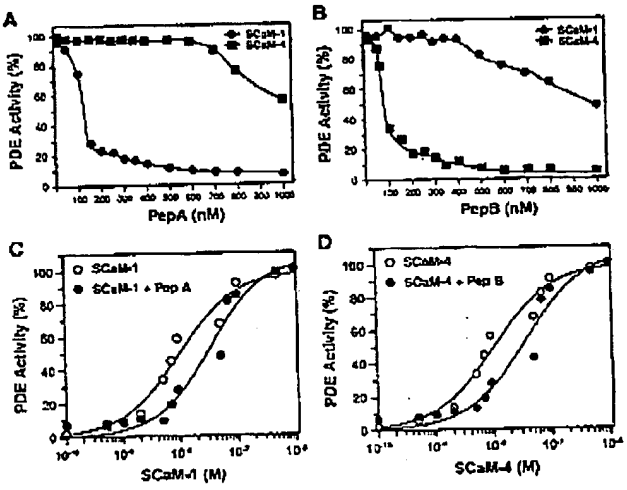


FIG. 9. **Effect of Pep A and Pep B on the activation of PDE by SCaM isoforms.** The dose-dependent inhibitions of PDE activity by Pep A and Pep B are shown. The inhibitory activities were measured in the presence of Ca^{2+} and a fixed concentration (120 nM) of SCaM-1 (*A*) or SCaM-4 (*B*) with increasing concentrations of Pep A or Pep B. PDE activity was measured in the presence of varying concentrations of SCaM-1 or SCaM-4 and either in the presence or absence of fixed concentrations (100 nM) of Pep A (*C*) or Pep B (*D*). Results are shown as the means obtained from three independent assays (n = 3).

and SCaM-4 (22). In Fig. 9, *A* and *B*, the effect of increasing peptide concentrations on PDE activity is shown for a fixed concentration (120 nM) of SCaM-1 or SCaM-4, a concentration sufficient for the maximal activation of PDE. A gradual decrease in PDE activity was observed with increasing concentrations of the relevant peptides. Half-maximal inhibition of PDE activation by SCaM-1 and SCaM-4 was obtained at ~120 nM Pep A and ~80 nM Pep B, respectively. Up to a concentration of 600 nM, Pep A had no effect on the activation of PDE by SCaM-4 and a concentration of more than 1 μM was required to reach half-maximal inhibition. Pep B, on the other hand, was a slightly more potent inhibitor of PDE activation by SCaM-1, inhibiting at concentrations above 400 nM, and although similar to Pep A, half-maximal inhibition was not reached at concentrations below 1 μM. To determine K_d values for these peptides in the activation of PDE by SCaMs, the CaM dose-dependent activation of PDE was determined in the presence (100 nM) or absence of the peptides (Fig. 9, *C* and *D*). The activation curves shifted to the right in the presence of the peptides, indicating a competition occurring between PDE and the peptides for the SCaM isoforms. The K_d values for Pep A and Pep B for their inhibition of the activation of PDE by SCaM-1 or SCaM-4 was determined to be 31.0 and 28.5 nM, respectively (Fig. 9, *C* and *D*). These experiments suggest that Pep A and Pep B have at least a 10-fold specificity for SCaM-1 and SCaM-4, respectively.

DISCUSSION

We have shown previously (22, 26–28, 33) that SCaM isoforms exhibit differences in their abilities to activate target enzymes. Importantly, some CaM isoforms competitively inhibit the activation of certain enzymes by other CaM isoforms, exhibiting a reciprocal regulation of target enzymes (26, 28). In these *in vitro* enzyme assays, the SCaM isoforms often possessed significantly different affinities for a given enzyme. For example, with the activation of plant Ca^{2+}-ATPase, SCaM-1 has ~20-fold higher affinity than SCaM-4 (28). Furthermore, in gel overlay assays of various plant CaM-binding proteins, these isoforms exhibited different binding intensities, suggesting that different CaM isoforms might require distinct primary sequences for optimal target binding (24). In this investigation, in which we took advantage of a phage-displayed random peptide library, we identify the peptide sequences that interact with specific SCaM isoforms. Surprisingly, two distinct families of peptides that specifically bind to SCaM-1 or SCaM-4 isoforms in a Ca^{2+}-dependent manner were identified in this study. These peptides fit into a 1-5-10 ((FILVW)*XXX*(FILVW)*XXXX*(FILVW)) motif for the SCaM-1-binding peptides and a 1-8-14 ((FILVW)*XXXXXX*(FAILVW)*XXXXX*(FILVW)) motif for the SCaM-4-binding peptides. Rhoads and Friedberg (17) originally identified these motifs as Ca^{2+}-dependent brain CaM-binding sequences, based upon the positions of conserved hydrophobic amino acid residues. These hydrophobic residues play an important role in anchoring peptides to target binding pockets of CaM via a hydrophobic interaction. It is also noteworthy that these Ca^{2+}-dependent CaM-binding motifs require basic residues, which are important in stabilizing the binding interaction by forming salt bridges. Consistent with this, all of the isolated peptides in this study were determined to be basic.

It is quite remarkable that plant CaM isoforms exhibit differing optimal sequence preferences for their target interactions *in vitro*. This is most likely due to structural differences between CaM isoforms. For example, SCaM-1 and SCaM-4 are ~22% different in regard to their primary amino acid sequences. What is the significance of this finding in terms of the function of CaM in plant cells? One intriguing possibility is that certain CaM isoforms may require specific binding targets thereby leading to unique cellular responses. This is consistent with the idea that these CaM isoforms evolved independently following segregation from a progenitor CaM, as predicted from phylogenetic analyses of these CaM isoforms (22). However, we should also point out that these isolated peptides could be the optimal binding sequences for given CaM isoforms, which are not found in nature. Indeed, in CaM gel overlay assays using plant cell extracts, we could hardly find binding proteins specific for the CaM isoforms (24). In addition, so far none of the enzymes tested *in vitro* in our laboratory exhibit specific binding to particular CaM isoforms. Furthermore, the differences in binding affinity between Pep A and Pep B to the SCaM isoforms were no greater than 10-fold. Therefore, it is reasonable to speculate that a large number of CaM-binding proteins bind to both isoforms, given a certain degree of affinity differences, and therefore only a few CaM isoform-specific binding proteins may exist in plants.

We reported previously that SCaM isoforms are different in their abilities to activate various target enzymes (28, 33). We

examined whether there is a correlation between the differential target binding preferences for the activation profiles of SCaMs. Interestingly, in the case of α-CaMKII, which has a 1-5-10 CaM-binding motif, the 1-5-10-favoring SCaM-1 activated this enzyme and at a much lower concentration than SCaM-4 (~10-fold less, K_{act} 22 *versus* 275 nM). Conversely, SCaM-4 was >4-fold better than SCaM-1 in the activation of nitric-oxide synthase, a 1-8-14-type enzyme (28). These positive correlations further support the biological relevance of our findings.

Previously, two other groups (39, 40) reported isolating CaM-binding sequences using similar strategies. Dedman *et al.* (39) pioneered this type of work using a random 15-mer library, shorter than the typical CaM-binding domains, to screen for CaM-binding peptides. Among the 28 isolated peptides, Trp was always present and often (11 peptides) present in the first variable position of the random peptide inserts. Additionally, 17 of the peptides contained Trp-Pro sequences. Nevalainen *et al.* (40) used a random 8-mer peptide library and similarly found the Trp residue and the Trp-Pro combination in the isolated CaM-binding peptides. Because they used a random peptide library consisting of lengths shorter than those of the naturally occurring CaM-binding sequences, the peptides binding to CaM might be the flanking amino acid sequences present in the vector, potentially posing a significant bias. In this regard, the 22-mer random peptide library used in this study has merit over those peptides possessing pre-determined constraints, and it is also long enough to span the length of the hydrophobic binding surface of CaM. These advantages might prove useful in the isolation of CaM-binding sequences that are more like naturally occurring CaM-binding peptides. The majority of the 30 plus CaM-binding domains identified thus far consists of stretches of 16–35 amino acid residues. Of the 33 peptides isolated in this study, Trp residues are found present at random positions (6 of them are actually devoid of the Trp residue), and only two of the peptides (*i.e.* α12 and β16) contain the Trp-Pro configuration. In addition, the 1-5-10 and 1-8-14 consensus motifs found in this study were originally identified in natural CaM-binding sequences, further arguing that the peptides isolated in this study resemble naturally occurring sequences.

What is the importance of our findings in regard to the structural perspective of the CaM-target peptide interaction? To address this, we built model structures of SCaM-1 bound to either M13 (of the 1-8-14 motif) or to the CaMKII CaM-binding sequence (of the 1-5-10 motif). The same models were prepared for SCaM-4. We then looked for any steric hindrance between SCaM and each of the peptides, hoping to discover any structural reason that would support our observation of the differential binding specificity. In short, we could not find any obvious reason that would clearly explain the basis of the binding discrimination of the SCaMs. However, we should note that this kind of modeling study assumes that both of the SCaM isoforms adopt the same conformation as that of target-bound mammalian CaM. Thus, the lack of a positive answer resulting from the model studies suggests two possibilities. First, SCaM isoforms may contain subtle differences in their side chain conformations of the residues directly interacting with target peptides despite the fact that their overall structures are similar to that of mammalian CaM. This would be similar to the structure of yeast CaM which is only 60% identical to mammalian CaM at the amino acid level (41). Indeed, we noticed that many Met residues in the SCaM isoforms were substituted for bulky hydrophobic amino acids such as Ile, Leu, and Val, which are involved in stabilizing the CaM-target peptide complex, by forming hydrophobic interactions and are also important for conferring flexibility to accommodate and recognize a wide range of targets (12). Alternatively, the overall structure of the SCaM isoforms might be different from that of mammalian CaM and may use different mechanisms for target binding than that of mammalian CaM. Therefore, only by high resolution three-dimensional structure determination of the complex with different peptides can the exact mechanism underlying the differential binding specificity of the SCaMs be determined.

The potential merits of this study are several fold. First, the isolated CaM-binding sequences may serve as good references for searching for CaM-binding proteins in data bases or for mapping the CaM-binding domains of certain proteins. Second, these peptides may serve as useful tools for studying CaM isoform-specific functions in plant cells. For example, in regard to each other, Pep A and Pep B exhibit at least a 10-fold higher specificity for SCaM-1 and SCaM-4, respectively, in both gel overlay and PDE inhibition assays. Another advantage of these techniques is that one can study the function of CaM in different subcellular compartments by tagging the inhibitor peptide to various subcellular targeting sequences. For example, its nuclear function can be studied by fusing it to a nuclear localization signal (6). Currently these studies are underway in our laboratory.

Acknowledgments—We are indebted to Dr. Mitsuhiko Ikura and Kyoko L. Yap (Department of Medical Biophysics, University of Toronto) for help in the modeling studies.

REFERENCES

1. Lu, K. P., and Means, A. R. (1993) *Endocr. Rev.* 14, 40–58
2. Crivici, A., and Ikura, M. (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 85–116
3. Bush, D. S. (1995) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46, 95–122
4. Snedden, W. A., and Fromm, H. (1998) *Trends Plant Sci.* 3, 299–304
5. Zielinski, R. E. (1998) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 697–725
6. Wang, J., Campos, B., Jamieson, G. A., Kaetzel, M. A., and Dedman, J. R. (1995) *J. Biol. Chem.* 270, 30245–30248
7. Lee, A., Wong, S. T., Gallagher, D., Li, B., Storm, D. R., Scheuer, T., and Catterall, W. A. (1999) *Nature* 399, 155–159
8. Liao, B., Paschal, B. M., and Luby-phelps, K. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 6217–6222
9. Zuhlke, R. D., Pitt, G. S., Deisseroth, K., Tsien, R. W., and Reuter, M. (1999) *Nature* 399, 159–162
10. Babu, Y. S., Bugg, C. E., and Cook, W. J. (1988) *J. Mol. Biol.* 204, 191–204
11. O'Neil, K. T., and DeGrado, W. F. (1990) *Trends Biochem. Sci.* 15, 59–64
12. Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science* 256, 632–638
13. Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) *Science* 257, 1251–1255
14. Meador, W. E., Means, A. R., and Quiocho, F. A. (1993) *Science* 262, 1718–1721
15. Afsha, M., Caves, L. S. D., Guimard, L., Hubbard, R. E., Calos, B., Grassy, G., and Haiech, J. (1994) *J. Mol. Biol.* 244, 554–571
16. Yuan, T., Vogel, H. J., Sutherland, C., and Walsh, M. P. (1998) *FEBS Lett.* 431, 210–214
17. Rhoads, A. R., and Friedberg, F. (1997) *FASEB J.* 11, 331–340
18. Yap, K. L., Kim, J., Truong, K., Sherman, M., Yuan, T., and Ikura, M. (2000) *J. Struct. Funct. Genomics* 1, 8–14
19. Yang, T., Segal, C., Abbo, S., Fedman, M., and Fromm, H. (1996) *Mol. Gen. Genet.* 252, 684–694
20. Poovaiah, B. W., Takezawa, D., An, G., and Han, T. I. (1996) *J. Plant Physiol.* 149, 553–558
21. Gawienowski, M. C., Szymanski, D., Perera, I. Y., and Zielinski, R. E. (1993) *Plant Mol. Biol.* 22, 215–225
22. Lee, S. H., Kim, J. C., Lee, M. S., Heo, W. D., Seo, H. Y., Yoon, H. W., Hong, J. C., Lee, S. Y., Bahk, J. D., Hwang, I., and Cho, M. J. (1995) *J. Biol. Chem.* 270, 21806–21812
23. Falke, J. J., Drake, S. K., Hazard, A. L., Peersen, D. B. (1994) *Q. Rev. Biophys.* 27, 219–290
24. Lee, S. H., Kim, M. C., Heo, W. D., Kim, J. C., Chung, W. S., Park, C. Y., Park, H. C., Cheong, Y. H., Kim, C. Y., Lee, S., Lee, K. J., Bahk, J. D., Lee, S. Y., and Cho, M. J. (1999) *Biochim. Biophys. Acta* 1433, 56–67
25. Heo, W. D., Lee, S. H., Kim, M. C., Kim, J. C., Chung, W. S., Chun, H. J., Lee, K. J., Park, C. Y., Park, H. C., Choi, J. Y., and Cho, M. J. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 766–771
26. Cho, M. J., Vaghy, P. L., Kondo, R., Lee, S. H., Davis, J. P., Rehl, R., Heo, W. D., and Johnson, J. D. (1998) *Biochemistry* 37, 15593–15597
27. Kondo, R., Tikunova, S. B., Cho, M. J., and Johnson, J. D. (1999) *J. Biol. Chem.* 274, 36213–36218
28. Lee, S. H., Johnson, J. D., Walsh, M. P., Van Lierop, J. E., Sutherland, C., Xu, A., Snedden, W. A., Kosk-Kosicka, D., Fromm, H., Narayanan, N., and Cho, M. J. (2000) *Biochem. J.* 350, 299–306
29. Chung, W. S., Lee, S. H., Kim, J. C., Heo, W. D., Kim, M. C., Park, C. Y., Park, H. C., Lim, C. O., Kim, W. B., Harper, J. F., and Cho, M. J. (2000) *Plant Cell* 12, 1393–1407

30. Pierce, H. H., Schachat, F., Brandt, P. W., Lombardo, C. R., and Kay, B. K. (1998) *J. Biol. Chem.* 273, 23448–23453
31. Guan, K. L., and Dixon, J. E. (1991) *Anal. Biochem.* 192, 262–267
32. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd Ed., pp. 5.3–6.59, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
33. Lee, S. H., Seo, H. Y., Kim, J. C., Heo, W. D., Chung, W. S., Lee, K. J., Kim, M. C., Cheong, Y. H., Choi, J. Y., Lim, C. O., and Cho, M. J. (1997) *J. Biol. Chem.* 272, 9252–9259
34. Bohm, G., Muhr, R., and Jaenicke, R. (1992) *Protein Eng.* 5, 191–195
35. Erickson-Vittanen, S., and Degrado, W. F. (1987) *Methods Enzymol.* 139, 455–458
36. Vogel, H. J. (1994) *Biochem. Cell Biol.* 72, 357–376
37. Yuan, T., Weljie, A. M., and Vogel, H. J. (1998) *Biochemistry* 37, 3187–3195
38. Keller, D., Clausen, R., Josefsen, K., and Led, J. J. (2001) *Biochemistry* 40, 10732–10740
39. Dedman, J. R., Kaetzel, M. A., Chan, H. C., Nelson, D. J., and Jamieson, G. A. (1993) *J. Biol. Chem.* 268, 23025–23030
40. Novalainen, L. T., Aoyama, T., Ikura, M., Crivici, A., Yan, H., Chun, N. H., and Nairn, A. C. (1997) *Biochem. J.* 321, 107–115
41. Ishida, H., Takahashi, K., Nakashima, K., Kumaki, Y., Nakata, M., Hikichi, K., and Yazawa, M. (2000) *Biochemistry* 39, 13660–13668

Biochem. J. (1999) 343, 29–37 (Printed in Great Britain)

Enzymic characterization *in vitro* of recombinant proprotein convertase PC4

Ajoy BASAK*†‡[1], Bakary B. TOURÉ§, Claude LAZURE‡, Majambu MBIKAY*, Michel CHRÉTIEN*† and Nabil G. SEIDAH§

*Molecular Neuroendocrinology Laboratory, Clinical Research Institute of Montréal, Montréal, Qué, Canada H2W 1R7,
†Protein Chemistry Center, Loeb Health Research Institute, Ottawa Hospital Medical Hospital, 725 Parkdale Avenue, Ottawa, ON, Canada K1Y 4K9,
‡Biochemical Neuroendocrinology Laboratory, Clinical Research Institute of Montréal, Montréal, Qué, Canada H2W 1R7, and
§Structure and Metabolism of Neuropeptides Laboratory, Clinical Research Institute of Montréal, Montréal, Qué, Canada H2W 1R7

Proprotein convertase PC4A, a member of the subtilisin/kexin family of serine proteases, was obtained in enzymically active form following expression of vaccinia virus recombinant rat (r)PC4A in GH4C1 cells. It displayed maximal activity at pH 7.0 and a Ca^{2+} concentration of 2.0 mM. Using PC4-specific antibodies, Western blot analysis of the medium revealed a major band at ≈ 54 kDa, corresponding to the molecular size of mature rPC4A. Among the various peptidyl-[4-methylcoumarin 7-amide (MCA)] substrates tested, the one that was preferred the most by rPC4A was acetyl (Ac)-Arg-Lys-Lys-Arg-MCA, which is cleaved 9 times faster (as judged from V_{max}/K_m measurements) than the best furin and PC1 substrate, pGlu-Arg-Thr-Lys-Arg-MCA. Recombinant rPC4A, along with human (h)furin and hPC1, cleaved a 17-amino-acid synthetic peptide, YQT**LRRRVKR**↓SLVVPTD (where ↓ denotes site of cleavage, and the important basic residues are shown in bold), encompassing the junction between the putative pro-segment of rPC4A and the active enzyme, suggesting a possible auto-activation of the enzyme. In an effort to identify potential physiological substrates for PC4, studies were performed with pro-[insulin-growth-factor (IGF)]-derived synthetic peptides, namely Ac-PAKSAR↓SVRA (IGF-I^{66-75}) and Ac-PAKSER↓DVST (IGF-II^{62-72}), as well as two lysine mutants [(IGF-$I^{66-75}Lys^{70}$) and (IGF-$II^{62-72}Lys^{67}$)]. Unlike PC1 and furin, rPC4A cleaved efficiently both IGF-I^{66-75} and IGF-II^{62-72}, suggesting a possible role of PC4 in the maturation of IGF-I and -II. In contrast, the peptides with a position 2 (P2) lysine mutation, IGF-$I^{66-75}Lys^{70}$ and IGF-$II^{62-72}Lys^{67}$, were cleaved more efficiently by PC1 and furin compared with rPC4A. Furthermore, using synthetic peptides containing the processing sites of pituitary adenylate-cyclase-activating polypeptide (PACAP)-38, we were able to confirm that, of the two testicular enzymes PC4 and PC7, PC4 is the best candidate enzyme for maturation of PACAP. Our data suggest that rPC4A is a functionally active convertase, with a substrate specificity somewhat different from that of other convertases, namely KXXR↓ (where X denotes any other residue). As expected, *p*-chloromercuribenzoic acid and metal chelators such as EDTA, EGTA and *trans*-1,2-diaminocyclohexane-*N*,*N*,*N*′,*N*′-tetraacetic acid inhibit the proteolytic activity of rPC4A, whereas it is activated by dithiothreitol. PC4A was also inhibited by transition-metal ions ($Cu^{2+} > Hg^{2+} > Zn^{2+}$ $Ni^{2+} > Co^{2+}$), as well as by small peptide semicarbazones (SCs), such as Arg-Lys-Lys-Arg-SC (K_i 0.75 μM) and Arg-Ser-Lys-Arg-SC (K_i 11.4 μM).

Key words: inhibitor, protease activity, protein substrate, substrate specificity.

INTRODUCTION

Following the discovery of proprotein convertases (PCs), the mammalian counterparts of the serine proteases of yeast kexin and bacterial subtilisin, it became important to study their structural and functional properties. Consistent with their substrate specificity, these convertases share a common function of processing a wide variety of pro-hormones and/or pro-proteins, as established by both *in vivo* and *in vitro* studies. These include neuropeptides, biologically active proteins and their receptors, growth factors, viral glycoproteins, surface proteins, and others [1,2]. In all cases, cleavage occurs on the C-terminal side of either a single Arg or a pair of basic residues (mainly Lys-Arg or Arg-Arg). To date, seven members of this family have been discovered. These are furin [also called paired basic amino acid converting enzyme (PACE)], PC1 (also called PC3), PC2, PACE4, PC4, PC5 (also known as PC6) and the most recently discovered member, PC7, also called LPC, PC8 or SPC7 [1,2]. The analysis of both tissue and cellular distribution, as well as gene regulation of the convertases, demonstrated a complex, but unique, pattern for each enzyme. Thus PC1 and PC2 are only expressed in endocrine and neuroendocrine tissues and cells. PC4 is found exclusively in germ cells, suggesting a possible reproductive function for this enzyme. PC5 and PACE4 are present in a wide array of tissue types [1,2]. Both furin and PC7 exhibit the most widespread tissue distribution. Varying degrees of overlap of proteolytic activity of convertases have been reported, indicating both redundancy and specificity in their biological functions [3–5].

Whereas the biosynthetic forms and *in vitro* properties of most convertases, including PC7 [6], have been examined quite extensively, little is known about PC4, which was discovered originally in 1992 [7,8]. In fact, the deduced cDNA sequence was the only criterion used for inclusion of this enzyme in the convertase family. Five mRNA transcripts (rPC4A–rPC4E) were identified, which differ with respect to their C- or N-terminal truncations [8,9]. It has been demonstrated that alternative

Abbreviations used: ac, acetyl; AMC, 7-amino-4-methylcoumarin; Boc, *t*-butoxycarbonyl; CDTA, *trans*-1,2-diaminocyclohexane-*N*,*N*,*N*′,*N*′-tetra-acetic acid; ECL, enhanced chemiluminescence; FAB, fast atom bombardment; h, human; m, mouse; MCA, 4-methylcoumarin 7-amide; PACAP, pituitary adenylate-cyclase-activating polypeptide; PACE, paired basic amino acid converting enzyme; PC(s), proprotein convertase(s); Pmc, 2,2,5,7,8-pentamethylchroman-6-sulphonyl; pro-IGF, proinsulin-like growth factor; r, rat; RP, reverse phase; SC, semicarbazone; VV, vaccinia virus.

[1] To whom correspondence should be addressed, at the Ottawa Hospital Medical School (e-mail abasak@lri.ca).

splicing is responsible for the generation of these 5′ and 3′ splice variants. The expression of mRNA for PC4 is largely restricted to testicular germ cells and, to some extent, within hormonally stimulated ovaries [7,8,10], whereas Western blot analysis revealed an immunoreactive protein in epididymal spermatozoa. The mouse (m)PC4 gene possesses an organization very similar to that of other convertases [8,11]. Recently, a PC4-null mouse has been obtained, and it was demonstrated that the fertility of homozygous null-animals is severely impaired [12]. It was also suggested that PC4 might be more important in the male than in the female for achieving fertilization, and for supporting early embryonic development in mice [12]. From their cDNAs, it was predicted that mPC4A and rat (r)PC4A would be soluble enzymes composed of 654 or 700 amino acids respectively. Co-expression of PC4 and several potential substrates, such as proenkephalins, pro-opiomelanocortin, 7B2, prorenin, amyloid precursor protein and pro-von Willebrand factor, by us and others [4,13] has so far yielded negative results. Until now, only one possible physiological substrate for PC4, namely pituitary adenylate-cyclase-activating polypeptide (PACAP), has been identified [14]. However, on the basis of localization in the testis and other studies, it is also likely that proinsulin-like growth factors (pro-IGFs)-I and -II could represent other potential substrates of PC4.

In the present study we report the first *in vitro* enzymic properties of recombinant rPC4A overexpressed in mammalian GH4C1 cells using the vaccinia virus (VV) expression systems. The processing of a large variety of synthetic fluorogenic substrates, as well as small peptides derived from pro-IGF-I, proIGF-II, PACAP and the junction of the pro-segment of PC4 and its catalytic subunit, were examined in order to elucidate the cleavage specificity and to identify possible substrates of PC4.

EXPERIMENTAL

Preparation of pCITE vector

The pCITE vector (Novagen; Madison, WI, U.S.A.) was digested with *NcoI*, blunted using the large fragment of DNA polymerase I (Klenow), and then digested with *XbaI* to generate a 3.6 kb band. All restriction and modification enzymes were purchased from Gibco Life Technologies (Burlington, ON, Canada).

Preparation of the 5′-end of rPC4

Two oligonucleotides (sense, 5′-GGCCCTCCCAGACAGCG-3′; antisense, 5′-AGATCTTGTTCTATCTCCTTGTTC-3′) of the rat PC4 sequence were used to amplify a 386-bp cDNA fragment by PCR. A typical PCR reaction consisted of a rPC4 cDNA plasmid as template at a concentration of 5 ng, 1 μM of both sense and antisense oligonucleotides, 2.5 mM $MgCl_2$, 1 × PCR buffer, 0.4 mM dNTPs and 0.5 unit of DNA *Taq* polymerase enzyme. The sense oligonucleotide was adapted with a linker (ATTACCATGG-) at its 5′-end, which was used to provide the initial methionine once the PC4 fragments were inserted into the proper open reading frame of the vector. The PCR product thus obtained was then blunted with T4 DNA polymerase, in the absence of dNTPs, to remove the 3′ adenosine generated by the PCR reaction. A *HinfI* digestion of the blunted PCR product yielded the 5′ rPC4 fragment of 236 bp used in the final construct. The 5′-end of PC4 is quite unusual, in that it is relatively short (32 nt), and is almost exclusively made up of G and C bases [9]. Its replacement as described above significantly improved PC4 mRNA translation *in vitro* (unpublished results), leading us to propose that this might increase in infected cells too.

Preparation of the 3′-end of rPC4

The 3′-end of rPC4 used in this construct came from an existing construction of rPC4 cDNA in the pBluescript vector [8], which was digested with *HinfI* and *XbaI* to yield a 2.2 kb band.

All the fragments used to generate this 5′-modified rPC4 construct were separated by agarose gel electrophoresis; the DNA bands of interest were excised and purified from the agarose using the Geneclean purification technique (BIO 101 Inc., La Jolla, CA, U.S.A.). The rPC4 5′- and 3′-fragments were ligated together in a 10:1 molar ratio, and the ligation reaction was confirmed by gel electrophoresis. Ligation of the blunt ends was used to generate *SspI* sites; the ligation reaction mixture was then digested with *SspI* (blunt end) and *XbaI* restriction enzymes to yield the 5′-modified rPC4 fragment of 2.4 kb. The pCITE-digested vector and the 5′-modified rPC4 fragment were then ligated together in a 1:3 molar ratio, and were used to transform XL1-Blue competent cells (Invitrogen, Carlsbad, CA, U.S.A.). Individual clones were analysed by restriction-enzyme digestion, and sequenced to confirm the accuracy of the reading frame [15–17].

Cellular infections and isolation of recombinant rPC4A

Essentially, a procedure similar to that reported previously for PC1 [18] was followed. In brief, GH4C1 cells, a mammalian somatomammotrophic cell line (Biosystems, CA, U.S.A.), were grown to confluence in air/CO_2 (9:1) at 37 °C in 15-cm-diam. tissue-culture dishes, using minimum essential medium supplemented with 10% (v/v) fetal bovine serum (Life Technologies Inc.) and 28 μg/ml gentamycin. Following a brief rinse with PBS, the cells were infected with recombinant viruses at 2–4 plaque-forming units/cell in 5 ml of PBS containing 0.01% (v/v) BSA for 30 min at room temperature. They were then returned to minimal essential medium plus fetal bovine serum, and incubated overnight at 37 °C. Cells were harvested twice, each time 3 h post-infection, by scraping into 20 mM Tris/acetate, pH 7.0. Suspensions from the Dounce homogenizer were centrifuged at 16000 *g* for 15 min at 4 °C in a microcentrifuge; the supernatants containing soluble cellular material were then removed and stored at −20 °C before further use. After collection of 2.5 l of medium, the frozen pools were thawed and concentrated to 3 ml (≈ 840-fold) on a Centriprep-30 (Amicon, Danvers, MA, U.S.A.) cartridge. This was then centrifuged at 30000 *g* for 15 min to eliminate all the suspended cellular particles, and the clear supernatant was separated into 100-μl aliquots that were stored at −20 °C. As a control, VV:WT/GH4C1 medium lacking rPC4A cDNA was similarly prepared, and stored for the purposes of comparison.

Antibodies against PC4

Three polyclonal antibodies raised against PC4, developed in the rabbit, were used in the present study (Figure 1). Antibody Ab-1 was raised against the synthetic peptide (D-Y^{74} + $rPC4^{75-90}$) of amino acid sequence D-Y^{74}QTLRRRVKRSLVVPTD^{90} located in the N-terminal domain of rPC4A [8]; antibody Ab-2 was developed using a C-terminal fusion protein ($rPC4A^{351-624}$); and antibody Ab-3 was generated against the whole PC4A protein by injecting the plasmid containing PC4A cDNA, following the procedure reported previously [19].

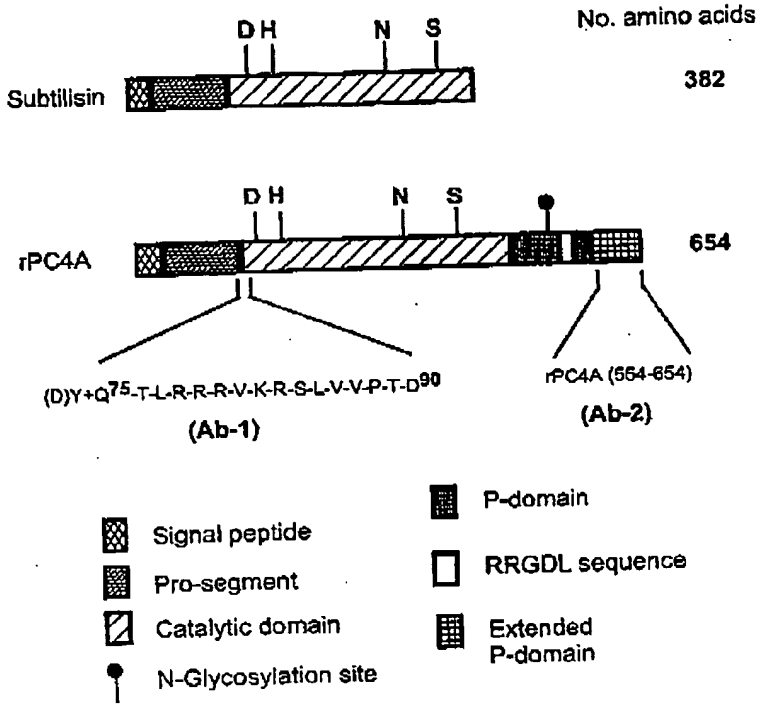


Figure 1 Schematic representation of the structure of rPC4A and subtilisin BPN'

The various domains of these proteases are depicted according to the key shown at the bottom of the Figure. The total number of amino acids featured in each enzyme is indicated on the right. The various amino acid regions used to produce rPC4A-specific antisera are also shown.

Western blot analysis

Media from rPC4A-infected GH4C1 cells were analysed by using SDS/PAGE (10% gels), and the separated proteins were electro-transferred on to PVDF membranes (Schleicher and Schuell, Keene, NH, U.S.A.). Protein bands were revealed by enhanced chemiluminescence (ECL) (Boehringer Mannheim, Indianapolis, IN, U.S.A.) using a primary rabbit antiserum [20] against either Ab-1 or Ab-2 (Figure 1). The specificity of the antiserum was verified by demonstrating that the bands made visible in the rPC4A-infected cells or medium could be blocked by pre-incubation with 100 μg/ml recombinant rPC4A protein.

Radioimmunoassays

For quantitative estimation of immunoreactive rPC4A protein in the concentrated medium and in cells obtained from VV:rPC4A-infected GH4C1 cells, three specific polyclonal antibodies, Ab-1, Ab-2 and Ab-3, were used in a procedure described previously [12,15]. Three separate samples of aliquots of various sizes were used, and the values were averaged.

Enzymic assay

The proteolytic activity of recombinant rPC4A was assayed using the commercially available synthetic peptidyl fluorogenic substrate, pGlu-Arg-Thr-Lys-Arg-[4-methylcoumarin-7-amide (MCA)] (Peptides International, Louisville, KY, U.S.A.) [18] as described below (hereafter, the residues are referred to by their single-letter codes). The enzyme sample (usually 10 μl) was routinely added to the assay mixture of a ternary buffer system consisting of sodium acetate, Mes and *N*-ethylmorpholine (all at 30 mM final concentration) and 2 mM $CaCl_2$, pH 7.0. The reaction (in a final vol. of 100 μl) was initiated at room temperature in a microtitre-well plate by the addition of pERTKR-MCA at a final concentration of 100 μM. Fluorescence readings were recorded after 60 min of incubation using a Perkin–Elmer spectrofluorimeter (model LS 50B), with the excitation and emission wavelengths being set at 370±3 and 460±5 nm respectively.

pH, Ca^{2+} and inhibitor profiles

The protocols used were essentially the same as reported elsewhere [6,18]. For pH-sensitivity experiments, enzymic activity was measured using the previously described ternary buffer system, with the pH being adjusted in the range from 4.0 to 8.0 with 0.5 pH-unit increments by using either acetic acid or NaOH. In order to estimate the Ca^{2+} concentration required for optimal activity, increasing concentrations of Na_2EDTA were added initially to the assay mixture as described above in the absence of added $CaCl_2$ until no more enzymic activity could be detected. On the basis of this result, the assay mixture was supplemented with 1 mM Na_2EDTA and various concentrations of $CaCl_2$, ranging from 0–5 mM (final). For inhibition studies, the enzyme sample was preincubated with selected compounds in the above assay mixture for 30 min prior to the addition of pERTKR-MCA. The solutions of inhibitors were prepared in solvents as follows: EDTA, EGTA, *trans*-1,2-diaminocyclohexane-*N*,*N*,*N*',*N*'-tetraacetic acid (CDTA), *p*-phenylazophenylarsonic acid and 4-chloromercuribenzoic acid were prepared in 1.25 M

NaOH; PMSF, 4-amidino-PMSF and 3,4-dichloroisocoumarin were prepared in 2% (v/v) DMSO; and pepstatin was dissolved in 10% (v/v) methanol. The rest of the inhibitors used were dissolved in water at pH 6.8. For experiments involving EDTA, EGTA and CDTA, no $CaCl_2$ was added to the assay medium.

Fluorogenic substrates and determination of K_m (app) and V_{max} values

pERTKR-MCA was substituted with various other peptidyl-MCA substrates either obtained from commercial sources or made synthetically in the laboratory [21]. For the determination of K_m values, increasing amounts of peptidyl-MCA were added to the reaction mixture. The data, plotted as the rate of hydrolysis against concentration of peptidyl-MCA used, were subjected to non-linear regression analysis using the Enzfitter software (Elsevier-Biosoft, Cambridge, U.K.) to determine the K_m and V_{max} values.

Synthesis of peptide substrates

The following peptides were selected: residues 66–75 from pro-IGF-I [acetyl (Ac)-PAKSAR↓SVRA] (Ia), residues 66–75 from pro-IGF-I, with an $Ala^{70} \rightarrow Lys$ substitution (Ac-PAKSKR↓SVRA) (Ib), residues 63–72 from pro-IGF-II (Ac-PAKSER↓DVST) (IIa), residues 63–72 from pro-IGF-II, with a $Glu^{67} \rightarrow Lys$ substitution (Ac-PAKSKR↓DVST) (IIb), D-Tyr+residues 75–90 from pro-rPC4A (D-YQTLRRRVKR↓SLVVPTD) (III), residues 170–176 from PACAP (GRR↓IPYL) (IVa), residues 168–176 from PACAP (NKGRR↓IPYL) (IVb), residues 165–176 from PACAP (RVKNKGRR↓IPYL) (IVc), residues 128–135 from PACAP (LSKR↓HSDG) (Va) and residues 122–135 from PACAP (DDDSEPLSKR↓HSDG) (Vb). (Here, and subsequently throughout the paper, ↓ denotes the site of cleavage, and the important basic residues are shown in bold.)

The synthesis of these peptides was performed on an automated solid-phase peptide synthesizer (model 431A; Applied Biosystems, Foster City, CA, U.S.A.) using 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate-mediated Fastmoc chemistry [22,23]. All amino acids (L-configuration) were protected as fluoren-9-ylmethoxycarbonyl (Fmoc) derivatives at the N-terminal. The coupling agents and solvents were purchased from Calbiochem–Novabiochem (La Jolla, CA, U.S.A.), Bachem (Torrance, CA, U.S.A.) and PE-Biosystems (Foster City, CA, U.S.A.). The following side chain-protecting groups were used: t-butyloxycarbonyl (Boc) for lysine; t-butyl for serine, threonine, tyrosine, aspartatic and glutamatic acids; and 2,2,5,7,8-pentamethylchroman 6-sulphonyl for arginine. At the end of the synthesis, the peptides were cleaved from the resin and fully deprotected by treatment twice with reagent K at room temperature (5.0 ml) for 4 h each time [22,23]. The crude material obtained was purified by reverse phase (RP)-HPLC (vista 5500; Varian, Basel, Switzerland), which was performed on a C_{18} semi-preparative column (1.0 cm × 25 cm; Chromatographic Specialties Inc., St. Laurent, PQ, Canada) using, as the buffer system, an aqueous 0.1% (v/v) trifluoroacetic acid solution and an organic phase of CH_3CN, also containing 0.1% (v/v) trifluoroacetic acid. Elution was performed using a linear gradient from 5–60% of the organic phase in 60 min following a 5-min isocratic step at a flow rate of 2 ml/min, and the appearance of the eluate was monitored at a wavelength of 225 nm. The peptides were fully characterized by both fast atom bombardment (FAB)-MS and amino acid analysis. Amino acid analyses were performed following 18 h of hydrolysis in 5.7 M HCl at 110 °C in a sealed tube under a vacuum, and the hydrolysates were analysed using a modified Beckman 120C autoanalyser equipped with a Varian DS 604 integrator/plotter, followed by post-column ninhydrin detection.

Incubation of synthetic peptides with various convertases

Incubations of various pro-IGF and pro-rPC4-related peptides (Ia, Ib, IIa, IIb and III) were carried out in a total vol. of 500 μl containing 50 μg of the peptide for time intervals ranging from 1 h to 24 h with hPC1, hfurin and rPC4A, representing activities of 15–16, 12–13 or 5–6 nmol of free 7-amino-4-methylcoumarin (AMC) released/h at 25 °C respectively. The final incubation mixtures contained, respectively, the ternary buffer and 5.0 mM $CaCl_2$, pH 6.0, for hPC1, 1.0 mM $CaCl_2$, pH 7.0, for furin and PC7, and 2 mM $CaCl_2$, pH 7.0, for rPC4A. The PACAP-based synthetic peptides (IVa–c and Va and Vb) were similarly incubated with only rPC4A and hPC7 [4].

Following the incubation, the aliquots were treated with 200 μl of 0.1M Na_2EDTA to terminate the reaction, and then were analysed on a Vydac 218TP54 C_{18} analytical column (0.1 cm × 25 cm) according to the conditions described above, except that the flow rate was maintained at 1.0 ml/min. The detected peptides were collected manually and subsequently analysed for amino acid composition, as well as by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) or ion-spray MS.

Synthesis of peptide semicarbazones (SCs)

The two tetrapeptide SCs, RSKR-SC and RKKR-SC, were synthesized as described previously [22]. The K_i values for these peptides against rPC4A were determined using Dixon's plot analysis [22,23].

Inhibition of rPC4A activity by various transition-metal ions

Five bivalent metal ions, mostly of the transition-metal series (namely Cu^{2+}, Co^{2+}, Zn^{2+}, Hg^{2+}, and Ni^{2+}) were examined for their ability to inhibit rPC4A activity. Freshly prepared aqueous solutions of $CuCl_2$, $CoCl_2$, $ZnSO_4$, $Hg(OAc)_2$ and $NiCl_2$ were preincubated with rPC4A for 5 min at room temperature in an assay mixture (100 μl) consisting of 30 mM each of Mes, sodium acetate and *N*-ethylmorpholine plus $CaCl_2$ (2 mM final concentration), pH 7.0, following which pERTKR-MCA (100 μM final concentration) was added. After another 4 h of incubation, the release of free AMC was monitored in a fluorimeter. Various concentrations of metal ions ranging from 50 to 1000 μM were used in the present study, and the residual activity was compared with the control carried out in parallel in the absence of any metal ions.

RESULTS

In vitro enzymic activity of rPC4A

Recombinant rPC4A, obtained from the medium of VV:rPC4A-infected GH4C1 cells upon concentration and dialysis, displayed a varying degree of enzymic activity towards fluorogenic peptidyl-MCA substrates such as benzyloxycarbonyl (Z)-ARR-MCA, Z-AKK-MCA, Z-KKR-MCA, Ac-KSKR-MCA, Ac-RKKR-MCA, ornithine-SKR-MCA, Ac-RSKR-MCA, Z-REKR-MCA, Boc-RVRR-MCA, pERTKR-MCA and Ac-YEKERSKR-MCA. Among them, Ac-RKKR-MCA was found to be the most potent, being cleaved at least 6 times more efficiently than Ac-RSKR-MCA or pERTKR-MCA, the commonly used fluorogenic substrates of most convertases [18,22,24,25]. In fact V_{max}/K_m values for the hydrolysis of Ac-RKKR-MCA are 10.3- and 8.4-fold greater than those for Ac-RSKR-

Table 1 Kinetic constants for the hydrolysis of synthetic peptidyl-MCA substrates by rPC4A

MCA substrate	$V_{max.}$ ($10^{-6} \cdot M \cdot s^{-1}$)	K_m (app) ($10^{-3} \cdot M$)	$V_{max.}/K_m$ ($10^{-3} \cdot s^{-1}$)	Fold
pERTKR-MCA	0.42 ± 0.05	4.21 ± 0.1	0.099	0.15
Ac-RKKR-MCA	0.12 ± 0.01	0.15 ± 0.02	0.83	1.0
Ac-RSKR-MCA	0.53 ± 0.09	6.6 ± 0.2	0.08	0.1

MCA and pERTKR-MCA respectively (Table 1). Surprisingly, even a tripeptidyl-MCA such as Z-ARR-MCA, lacking a peptidyl residue at the P4 position, was found to be as good a substrate as RSKR-MCA, REKR-MCA and RVRR-MCA, all of which possess an additional P4 arginine residue. This observation is somewhat different from that reported for other PCs, and suggests that PC4 is a distinct and functionally active convertase for which the P3 position may be more critical than for other convertases. This study also demonstrates the first *in vitro* enzymic activity of PC4.

pH profile

Employing a buffer system of different pH values ranging from 4.0 to 8.0 (see the Experimental section), it was noted that the pH optimum of rPC4A resides in the vicinity of 7.0 (Figure 2), very similar to that observed for furin [26–28] and PC7 [6]. Interestingly, both PC7 and furin maintained a significantly greater percentage of their maximal activities, even at pH 5.5, in contrast with PC4, which lost more than 90 % of its maximal activity at this pH. Thus unlike PC7 and furin, PC4 displays a narrower pH optimum.

Ca^{2+} profile

A precise measurement of the Ca^{2+} requirement for maximal activity of rPC4A was accomplished, following prior measurement of endogenous Ca^{2+} in the enzyme preparation by titration with Na_2EDTA (results not shown). The Ca^{2+}-activation profile of rPC4A at pH 7.0 is shown in Figure 2, which suggests that the optimal Ca^{2+} concentration for rPC4A activity is near 2.0 mM, again reminiscent of results reported for furin [26–28] and PC7 [6].

Inhibitory profile

On the basis of the study using various general and class-specific inhibitors, it was noted that the enzymic activity of rPC4A, as expected, is highly sensitive to cationic chelators such as EDTA, EGTA and CDTA, all of which, at 1–10 mM concentration, abolished almost completely the full enzymic activity. These observations confirm that this activity is indeed Ca^{2+}-dependent. Furthermore, as found for other convertases, PC4A is also strongly inactivated by *p*-chloromercuribenzoate. Most cysteine protease and mixed serine/cysteine protease inhibitors had no significant effects on PC4A activity. *p*-Phenylazophenylarsonic acid, a known inhibitor of subtilisin [29], appears in this case to be a weak inhibitor of rPC4A, in contrast with its action towards PC1 and furin, both of which are strongly inhibited at a concentration of inhibitor of 0.1 mM (A. Basak and C. Lazure, unpublished work). Finally, a differential effect due to the presence of a reducing agent was observed: a 2–2.5-fold enhancement in PC4 activity occurred upon addition of 1 mM dithiothreitol. A similar effect was also noted for this compound on the activity of PC7 [6], although a completely opposite effect

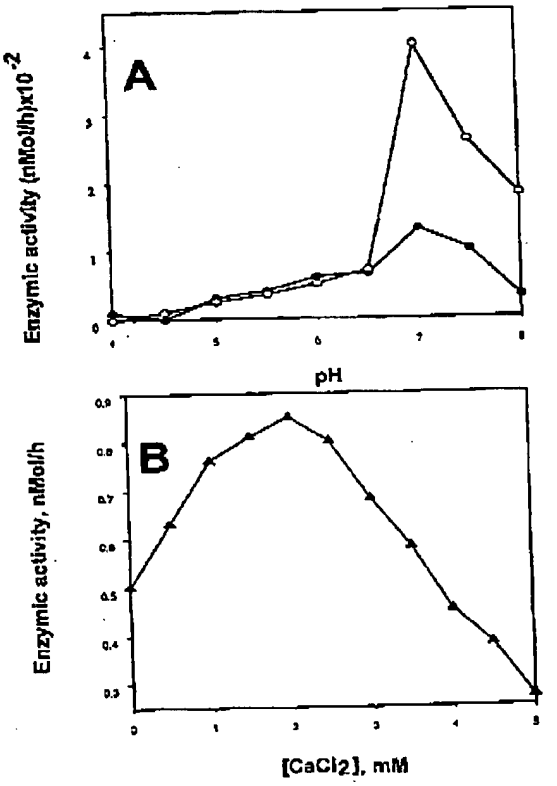


Figure 2 pH-activation (A) and Ca^{2+}-activation (B) profiles of recombinant rPC4A

(A) Preparations of VV:rPC4A-infected GH4C1 cells were assayed according to the protocol described in the Experimental section using pERTKR-MCA as a fluorogenic substrate. The results represent an average of three separate independent determinations for crude (●) and partially purified (○) rPC4A respectively. (B) Following prior incubation for 10 min with 1.0 mM EDTA to remove the endogenous Ca^{2+}, enzyme samples from the medium of VV:rPC4A-infected GH4C1 cell lines were assayed in the presence of increasing concentrations of $CaCl_2$, as described in the Experimental section.

was reported with other convertases, namely ykexin and soluble furin [6].

Inhibitory influence of bivalent transition-metal ions

Consistent with earlier observations with PC1, furin [30] and PC7 [11], a strong inhibitory effect on rPC4A activity by the transition-metal ions such as Cu^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+} and Hg^{2+} was also noted; this inactivation was found to occur (in terms of decreasing inhibitory effect) in the order of $Cu^{2+} > Hg^{2+} > Zn^{2+} > Ni^{2+} > Co^{2+}$, with IC_{50} measurements for Cu^{2+} and Co^{2+} of

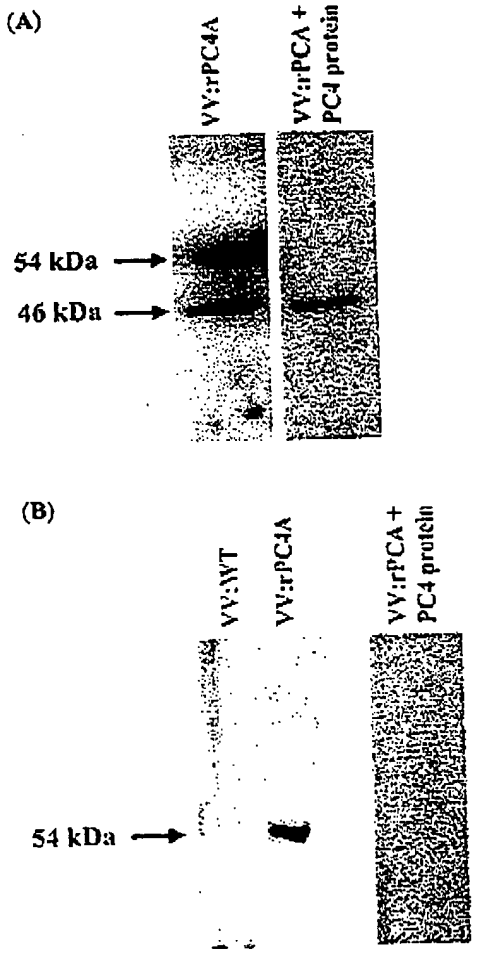


Figure 3 Immunodetection of rPC4A using Western blot analysis with Ab-1 antibody (A), or with Ab-2 antibody (B)

Samples from VV:rPC4A-infected GH4C1 cells were processed as described in the Experimental section, and analysed with SDS/PAGE (10% gels). Following electrotransfer to PVDF membranes, protein bands were made visible by ECL detection using a primary rabbit antiserum raised against the N-terminal peptide (Ab-1). (A) Left panel: VV:rPC4A-infected GH4C1 medium; right panel: VV:rPC4A-infected GH4C1 medium + rPC4A protein (100 μg/ml). Note that the band at 54 kDa, displaceable by PC4A protein, represents the full-length, active PC4A, whereas the band at 46 kDa is non-specific and not displaceable by PC4A protein. (B) The lanes were loaded as follows: left panel, first lane, VV:WT-GH4C1 medium; left panel, right lane, VV:PC4A-infected GH4C1 medium; and right panel, VV:rPC4A-infected GH4C1 medium + rPC4 protein (100 μg/ml).

12 ± 1.2 and 250 ± 4.5 μM respectively (data given as means $\pm$ S.D.). The possible cause for this inactivation may be related to interactions with metal-sensitive amino acid residues, such as cysteine, consequently leading to oxidation and/or complex formation [30].

Western blot analysis of rPC4A

Using the N-terminally directed antibody, Ab-1, Western blot analysis of the culture medium obtained from GH4C1 cells infected with VV:rPC4A revealed a distinct band at $\approx$ 54 kDa (Figure 3A). As expected, the appearance of this band can be fully abolished by the presence of recombinant rPC4A protein (100 μg/ml), thereby confirming its identity. A similar band, also displaceable by rPC4A protein (100 μg/ml), was noted in the cell medium using the C-terminally directed Ab-2 antibody (Figure 3B). This $\approx$ 54 kDa band represents the full-length form of mature rPC4A (544 amino acids), and the molecular mass compares reasonably well with that of the endogenous enzyme (M_r approx. 52000) that is expressed in germinal cells of rat testes, and which is absent in similar cells of PC4-null mice [12]. It may be pointed out that a large majority of the recombinant rPC4A protein remained associated with the cell membrane as insoluble material, which was only partially extractable upon repeated cell lysis (results not shown).

Radioimmunoassay

Radioimmunoassays carried out on various samples of recombinant rPCA using the antibodies Ab-1, Ab-2 and Ab-3 indicated the presence of an extremely low amount of immunoreactive PC4A. On average, this ranged from 10–14 μg/l of medium, which was more than 100-fold less than that obtained in the case of PC1 [18].

Digestion profiles of synthetic peptides

The distribution of mRNA or protein levels of biologically active protein/hormone precursors in testicular germ cells [8,9,12,14, 15,31] suggested a number of potential substrates of PC4A [12]. These include pro-nerve growth factor (THRSKR$\downarrow$SS), pro-IGF receptor (PERKRR$\downarrow$DV), (m)pro-IGF-I (PAKSAR$\downarrow$SV), (m)pro-IGF-II (PAKSER$\downarrow$DV), (m)pro-fertilin α-chain (QSRMRR$\downarrow$AA), (m)pro-fertilin β-chain (SCKLKR$\downarrow$RG), (m)ADAM5 (a disintegrin and metalloproteinase enzyme) (ARRPRR$\downarrow$IC), (h)proenkephalin (GGFMKR$\downarrow$HS), (r)pro-cyritestin (AERLCR$\downarrow$KS) and hPACAP-related peptide (AVLGKR$\downarrow$YK). In an effort to carry out chemical and enzymic characterizations, and also to determine potential substrates for PC4A, a number of synthetic peptides based on the sequences containing the putative cleavage site of pro-IGF-I (Ia and Ib), pro-IGF-II (IIa and IIb), pro-rPC4A (III) and PACAP (IVa–c and Va and Vb) were examined (Table 2). In addition to rPC4A, other convertases such as hPC1, hfurin and hPC7 were included in this study for comparison purposes.

As indicated in Table 2, both pro-IGF-I$^{69-75}$ (Ia) and pro-IGF-II$^{68-73}$ (IIa) were cleaved correctly by rPC4A at the expected physiological site, as confirmed by FAB-MS and amino acid analysis of the isolated products. In contrast, no significant processing of either (Ia) or (IIa) was noticed with hPC1 or hfurin, even after 24 h of incubation (Figure 4). As expected, the two lysine mutants [(Ib) and (IIb)] were both cleaved by PC4, PC1 and furin quite efficiently (Table 2), suggesting that, in the context of a monobasic cleavage, lysine at the P4 position within the sequence motif KXXR is recognized more easily by PC4 than by PC1 and furin.

Similarly, the pro-PC4-derived synthetic peptide (III) was also cleaved fairly efficiently by rPC4A, as well as by hfurin and hPC1 (Figure 4 and Table 2). The time required ($t_{1/2}$) for 50% cleavage of this peptide to take place by rPC4A, PC1 and furin, using an almost identical enzymic activity as measured against the fluorogenic substrate, pERTKR-MCA, showed that proPC4-derived peptide (III) was cleaved most efficiently by furin, followed by PC4 and PC1 respectively.

The synthetic PACAP-based peptides containing the N- and C-terminal processing sites (IVa, IVb, IVc and Va and Vb) were examined as possible substrates for both rPC4A and hPC7, since

Table 2 Relative cleavage rates of synthetic pro-PC4, pro-IGF-I, pro-IGF-II and PACAP-related peptides by various convertases, including rPC4A

The enzymic activity used in the digestion study of these synthetic peptides was 1.63, 1.38, and 1.45 nmol/h of released AMC from pERTKR-MCA (100 μM final concentration) for rPC4A, hPC1 and hfurin respectively. nd, not determined. For further details of the procedure used, see the Experimental section.

		Extent of cleavage (%)						
		VV:rPC4		VV:hPC1		VV:hfurin		VV:hPC7
Peptide sequence	Time ...	4 h	22 h	4 h	22 h	4 h	22 h	4 h
pro-IGF-I (66–75) (Ia) (Ac-PAKSAR ↓ SVRA)		39	71	7	10	0	0	
pro-IGF-I (66–75 Lys^{70}) (Ib) (Ac-PAKSKR ↓ SVRA)		nd	50	1	45	53	100	
pro-IGF-II (53–72) (IIa) (Ac-PAKSER ↓ DVST)		100	100	0	5	0	8	
pro-IGF-II (53–72 Lys^{67}) (IIb) (Ac-PAKSKR ↓ DVST)		69	81	3	73	69	100	
rPC4(D-Tyr + 75–90) (III) (D-YQTLRRRVKR ↓ SLVVPTD)		25	67	15	44	84	100	
PACAP (170–176) (IVa) (GRR ↓ TPYL)		< 5						< 5
PACAP (168–176) (IVb) (NKGRR ↓ TPYL)		78						< 5
PACAP (165–176) (IVc) (RVKNKGRR ↓ TPYL)		85						< 10
PACAP(128–135) (Va) (LSKR ↓ HSDG)		95						0
PACAP(122–135) (Vb) (DDDSEPLSKR ↓ HSDG)		90						2

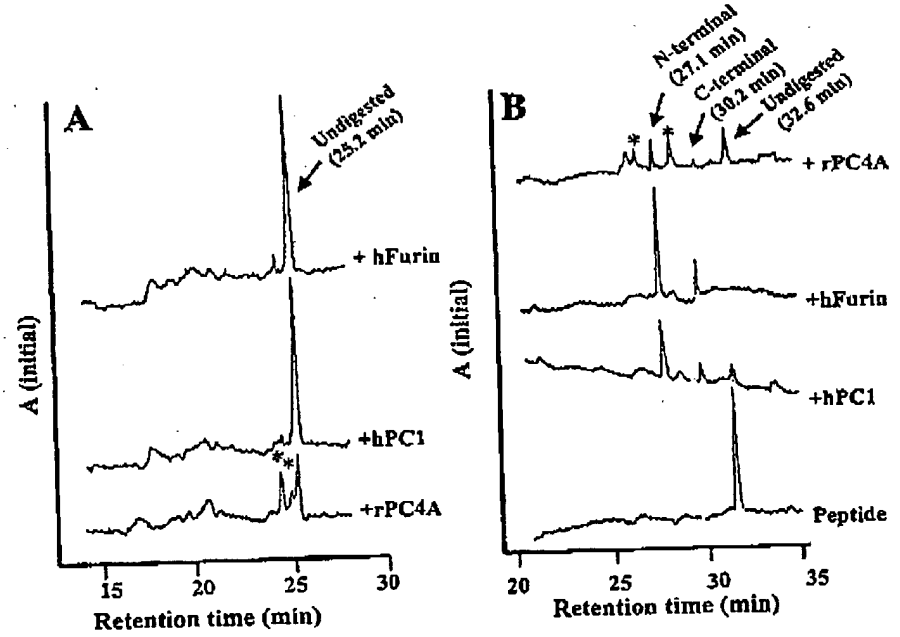


Figure 4 (A) Treatment of IGF-I peptide (Ia) [Ac-PAKSAR ↓ SVRA] with hPC1, hfurin and rPC4A, and (B) treatment of rPC4A peptide (III) (D-YQTLRRRVKR ↓ SLVVPTD) with rPC4A, hfurin and hPC1

(A) RP-HPLC analysis of the digestion products was performed as described in the Experimental section. For furin and PC1, the amount of peptide injected was twice the amount injected for PC4-digested material. The peaks in the chromatogram for PC4-digested peptide were characterized as follows: the earliest eluting peak (R_t = 23.9 min) was N-terminal peptide (Ac-PAKSAR), the peak eluting next (R_t = 24.7 min) was C-terminal peptide (SVRA) and the last peak to be eluted (R_t = 25.2 min) was undigested peptide (Ia). The newly generated peaks upon digestion are indicated by asterisks. (B) RP-HPLC analysis of the digestion reactions was carried out as described in the Experimental section. The peptide (III) (R_t = 32.6 min) was cleaved at the expected physiological site by rPC4A, hfurin and hPC1 to yield N-terminal peptide (D-YQTLRRRVKR) (R_t = 27.1 min) and C-terminal peptide (SLVVPTD) (R_t = 30.2 min). With PC4A, two additional peaks (indicated by asterisks) at R_t = 26.3 min and R_t = 28.0 min were observed. These were attributed to the cleaved peptides D-YQTLRR and RVKRSLVVPTD respectively, suggesting a second site of cleavage for PC4 as indicated (D-YQTLRR ↓ RVKRSLVVPTD).

these convertases were found to be co-expressed with PACAP in germ cells [14]. Our study revealed that both (IVb) and (IVc) were cleaved by PC4, but not by PC7, whereas (IVa) was not cleaved by either convertase, again reinforcing the result that, independently of other positions, the P4 lysine plays an important role in the substrate recognition by PC4. Again, both (Va) and (Vb) were cleaved fairly well by PC4, but not by PC7. These data *in vitro* suggest that PC4 is a better candidate enzyme for the

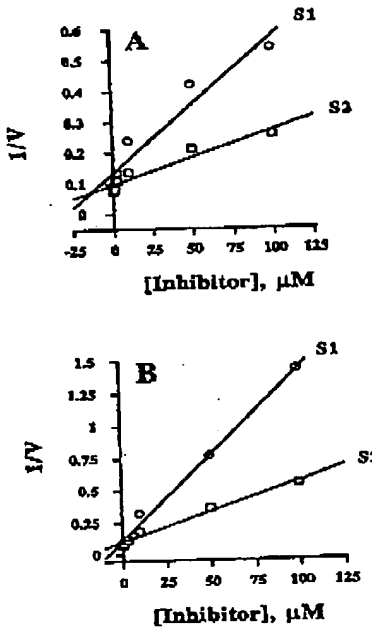


Figure 5 Dixon plot showing the inactivation of proteolytic activity of rPC4A by RSKR-SC (A) and RKKR-SC (B)

The activity of PC4A was followed at different substrate (pERTKR-MCA) concentrations for each inhibition study, namely 50 μM (S1) and 100 μM (S2).

processing of PACAP to PACAP-38, and possibly to PACAP-27, a notion already proposed by previous studies [14].

Ca^{2+}-dependency of rPC4A activity

The dependency of proteolytic activity of rPC4A on Ca^{2+} is demonstrated by the observation that PC4-mediated cleavage of pro-IGFs and pro-PC4-related synthetic peptides can be fully abolished by the addition of EDTA (50 μM) (results not shown).

Effect of RKKR-SC and RSKR-SC on rPC4A activity

Both tetrapeptides RKKR-SC and RSKR-SC exhibited a strong inhibition of rPC4A activity in a reversible, competitive manner. This was confirmed by Dixon plot analysis, as shown in Figure 5, which indicates further that the K_i values for these peptides are 0.75 and 11.4 μM respectively. Thus RKKR-SC is nearly 66-fold more potent as an inhibitor of rPC4A than is RSKR-SC. It may be pointed out that these peptidyl-SCs were shown to be fairly strong inhibitors of both PC1 and furin, with K_i values ranging from 1.4 to 15 μM [22].

DISCUSSION

Among all the members of convertases, PC4 is probably the most unique because of its exclusive distribution in testis and ovarian cells. On the basis of mRNA distribution in various cell lines and in rat tissues, it was observed that, within the germinal cells, PC4 is found in pachetene spermatocytes and spherical spermatids of the testis [7,8]. In mouse, expression is limited only to the round spermatids of germ cells [7]. These observations, and the latest study that shows that PC4 knock-out mice possess reduced fertility, sperm motility and are hyperactive compared with normal ones [12], suggest that PC4 probably plays an important role in reproductive biology, even though its mode of action has yet to be determined. Until now, very little was known about the enzymic properties of PC4 and its physiological substrates. The present study has now provided some important leads in this direction.

In vitro treatment of various synthetic peptides with rPC4A, hPC1 and hfurin confirmed that the peptide representing the junction between pro-PC4A and PC4A (YQTLRRRVKR↓SLVVPTD) was cleaved (as indicated) not only by rPC4A, but also by hPC1 and hfurin. This result indicated that, like other convertases, PC4 could also be autocatalytically activated. Both pro-IGF-I- and pro-IGF-II-derived peptides, Ac-PAKSAR↓SVRA (Ia) and Ac-PAKSER↓DVST (IIa), were cleaved at the site indicated by rPC4A, but not by PC1 and furin. It was further noted that the pro-IGF peptides that were mutated to lysine at the P2 position, namely Ac-PAKSKR↓SVRA (Ib) and Ac-PAKSKR↓DVST (IIb), were cleaved by all three convertases. These results suggest that a tetrapeptide sequence motif (KXXR) containing a P4 lysine and a P1 arginine residue is best recognized by rPC4A. The results indicate further that precursors of both IGF-I [32] and IGF-II [33,34] could represent potential substrates for PC4. It is known that these two important growth factors are generated from their precursors by the proteolytic removal of 35 and 38 amino acids respectively from their C-terminus (known as the E-domain), followed by the removal of Arg^{71} or Arg^{69} for IGF-I or IGF-II respectively by the action of carboxypeptidase. In both cases, there is a unique 13-residue prohormone cleavage motif, characterized by the presence of five basic residues: $KXXKXXR^{71/69}\downarrow XXRXXR$. Consistent with the PC4 distribution, IGF-I mRNA is highly expressed in germ cells, particularly after stimulation by follicle-stimulating hormone and, to a lesser extent, after treatment with growth hormone or luteinizing hormone [35]. Although IGF mRNA is largely localized specifically in Leydig cells, these cells are not the sole site of origin in the testis. In fact, non-Leydig cell expression might be an important component of testicular IGF-I production [36]. Numerous developmental abnormalities, such as growth deficiencies, delayed bone development, infertility and high mortality rate, in addition to many forms of cancer, have been linked to the expression of both IGF-I and -II [37–39]. Because of these important biological functions of IGF-I and -II, particular attention was taken to identify their processing enzymes. The studies performed to date [40,41] are sufficient to exclude furin as being the processing enzyme involved in IGF maturation, since furin-deficient Lovo cells were able to process pro-IGF-I at the correct $Arg^{71}\downarrow$ site, whereas transiently transfected HEK-293 cells process pro-IGF-I at the $Arg^{71}\downarrow$ site, as well as at another site C-terminal to this, possibly $Arg^{77}\downarrow$. In a recent study of co-expression employing constructs of furin, PACE4, PC5A, PC5B and PC7, Duguay et al. [41] demonstrated that PACE4 and PC5A/PC5B are the most likely candidates for being the pro-IGF convertase. Our studies have now added PC4A to the list as another candidate for being the processing enzyme of pro-IGF-I and pro-IGF-II in the testis, and have also confirmed the earlier conclusion that neither furin nor PC1 is able to produce mature IGF-I from its precursor.

Recently, PC4 has been implicated in the processing of PACAP [14]. PACAP was isolated from the hypothalamus, and is known to exist as two amidated forms with 38 or 27 amino acids (PACAP-38 or PACAP-27 respectively). Studies on the immunohistochemistry of rat testis demonstrated strong PACAP-like immunoreactivity in spermatids in both cap and acrosomal phases [41,42]. An almost identical pattern of expression of PC4